Maastricht University

Contextual interference

Citation for published version (APA):

Graser, J. V. (2021). *Contextual interference: the challenging way towards evidence in paediatric neurorehabilitation* . [Doctoral Thesis, Maastricht University]. Gildeprint Drukkerijen. https://doi.org/10.26481/dis.20210316jg

Document status and date:
Published: 01/01/2021

DOI:
10.26481/dis.20210316jg

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 19 Apr. 2024

CONTEXTUAL INTERFERENCE

THE CHALLENGING WAY TOWARDS EVIDENCE IN PAEDIATRIC NEURO REHABILITATION

JUDITH VERENA GRASER

© Copyright by Judith V Graser, 2021.
All rights reserved. No part of this thesis may be reproduced or transmitted in any form or any means, electronic or mechanical, including photography, recording or any information storage or retrieval system without permission in writing from the author, or, when appropriate, from the publisher of the publications.
Printed by: Gildeprint – www.gildeprint.nl
ISBN: 9789464191295

CONTEXTUAL INTERFERENCE – THE CHALLENGING WAY TOWARDS EVIDENCE IN PAEDIATRIC NEUROREHABILITATION

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

Tuesday, March 16 2021, at 16:00h

by
Judith Verena Graser

Supervisors
Prof. dr. R. A. de Bie
Dr. C. H. G. Bastiaenen

Co-Supervisor
Prof. dr. H. J. A. van Hedel, University Children's Hospital Zurich, Switzerland

Degree Committee
Prof. dr. F. J. M. Feron (chair)
Prof. dr R. J. Vermeulen
Prof. dr. A. Buizer VU University Medical Center, Amsterdam, the Netherlands
Dr. E. A. A. Rameckers
Dr. B. Tobler-Ammann, Insel University Hospital, Berne, Switzerland
Prof. dr. K. Klingels, Hasselt University, Belgium

The research presented in this dissertation was supported by the Clinical Research Priority Program, Neurorehabilitation, University of Zurich, Switzerland and the Mäxi Foundation, Zurich, Switzerland.

CONTENTS

CHAPTER 1	General introduction	9
CHAPTER 2	The role of the practice order: A systematic review about contextual interference in children	35
CHAPTER 3	Reliability of and practical lessons learnt from robotic upper limb assessments in children undergoing neurorehabilitation	79
CHAPTER 4	Contextual interference with ChARMin exergames in children with brain lesions and impaired upper limb function– a feasibility experiment	117
CHAPTER 5	Contextual interference in children with brain lesions: protocol of a pilot study investigating blocked vs. random practice order of an upper limb robotic exergame	139
CHAPTER 6	Contextual interference in children with brain lesions: a pilot study investigating blocked vs. random practice order of an upper limb robotic exergame	169
CHAPTER 7	General discussion	203
ADDENDUM		227

Summary
Zusammenfassung
Impact
Acknowledgement
About the author
List of publications

ABBREVIATIONS

95%CI	95% confidence interval
AACPDM	American Academy of Cerebral Palsy and Developmental Medicine
ACE	Absolute constant error
ANOVA	Analysis of variance
aROM	Active range of motion
ATD	Alternating treatment design
B	Significant (result), favouring blocked order
BBT	Box and Blocks test
CA	Chronological age
CCT	Controlled clinical trials
ChARMin	Children arm rehabilitation mechatronic interface
CI	Contextual interference
cm	Centimetre
CONSORT	Consolidated Standards of Reporting Trials
CP	Cerebral palsy
Diff	Difference
DS	Down's Syndrome
ES	Effect size
fMRI	Functional magnetic resonance imaging
FMT	Fine-motor task
GM	Grand mean
GMT	Gross motor task
IC	Inconsistent (results)
ICC	Intraclass correlation coefficient
ICF	International Classification of Functioning, Disability, and Health
ID	Identity
IDD	Intellectual developmental disability
kg	Kilogramme
LB	Lower bound
LD	Learning disabilities
LT	Laboratory task
m	Metre
M1	Median of the first measurement
M2	Median of the second measurement
m^3	Cubic metres
MA	Mental age
MA2	Melbourne Assessment 2

$MA2_{fluency}$	Melbourne Assessment 2, subscale fluency
$MA2_x$	Melbourne Assessment 2 at time point X
MACS	Manual Ability Classification System
MAS	Modified Ashworth scale
MB	Multiple baseline
MBD	Multiple baseline design
MDC	Minimal detectable difference
ms	Milliseconds
MSE	Mean square error
n	Number
NA	Not applicable
NLT	Non-laboratory task
Nm	Newton-metre
Nm/rad	Newton-metre per radian
nP_{norm}	Number of velocity peaks normalised to the distance of the covered distance of the movement path
NR	Not reported
NS	Not significant
OT	Occupational therapy
PT	Physiotherapy
PICO	Population, Intervention, Control, Outcome
pROM	Passive range of motion
QoM	Quality of movement
R	Significant (result), favouring random order
RCT	Randomised controlled trial
RLEX	Robotics lower extremities.
RUEX	Robotics upper extremities,
r	Pearson`s correlation coefficient
RPM	Resistance to passive movement
SD	Standard deviation
SD_{change}	Standard deviation of the change
SMS Lab	Sensory Motion Systems Lab
SORT	The Standardised Reporting of Trials
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials
SRD	Smallest real difference
SRD%	Smallest real difference divided by the grand mean x 100
SRM	Standardised response mean
SSRD	Single subject research design
t	Time point
TD	Typically developing
TMS	Transcranial magnetic stimulation

TONI-4	Test Of Nonverbal Intelligence – Version 4
UB	Upper bound
WSR test	Wilcoxon signed-rank test
X_{change}	Mean change score
yr./yrs.	Year/years
Z_{α}	Standard normal z-value for a significance level α=0.05
Z_{β}	Standard normal z-value for the power of 80%
α	Significance level
μ_1	Mean of intervention group 1
μ_2	Mean of intervention group 2
σ	Pooled standard deviation

CHAPTER 1

GENERAL INTRODUCTION

CHAPTER 1

"NOTHING HAPPENS UNTIL SOMETHING MOVES. […]."

Albert Einstein

Human beings are inherently social [1]. According to the Canadian psychologist Paul Bloom, we are happier and better, when connected to others [2]. Therefore, social behaviour is a central aspect of human life and ultimately what makes us human. Social interaction is closely connected to motor function. Basically, it is motor function that enables human beings to interact with others at all. When we talk, verbally or with body language, write, dance, kiss, or walk, every action essential to get and stay in connection with other human beings requires motor function.

MOTOR BEHAVIOUR

All activities occurring due to the stimulation of motor neurons can be gathered under the term "motor function", synonymously also called motor behaviour [3]. Motor behaviour addresses neural, physical and behavioural aspects of human movement [4]. It covers three distinct sub-disciplines: motor development, motor control and motor learning [4, 5]. These sub-disciplines are closely related to each other: "Motor control is the physiological process whereby motor development occurs, and motor learning allows motor development to occur systematically, resulting in a permanent change in motor behavior due to experience." p. 65 [6]. The definitions of motor behaviour and its three sub-disciplines motor development, motor control and motor learning, are displayed in Figure 1.1.

MOTOR DEVELOPMENT

Human motor function develops with the course of child development and starts at around week eight of pregnancy with startles and hick-ups [7]. Motor development is described as a process of individual maturation [7] during which new skills develop on the basis of previously developed skills. It is characterised by age-related systematic and constant changes lasting throughout the life span and the factors influencing them [4]. In the majority of human beings, motor development follows a general agenda of so called milestones, with individual deviations within the norms. Environmental factors influence this development. For example, a child having access to crayons at an early age, is more likely to hold it and draw with it earlier compared to a child who gets in contact with crayons at a later stage. Depending

on opportunities children encounter in their daily environment, the differentiation of the maturated motor function is further developed [7] throughout the whole life span [8].

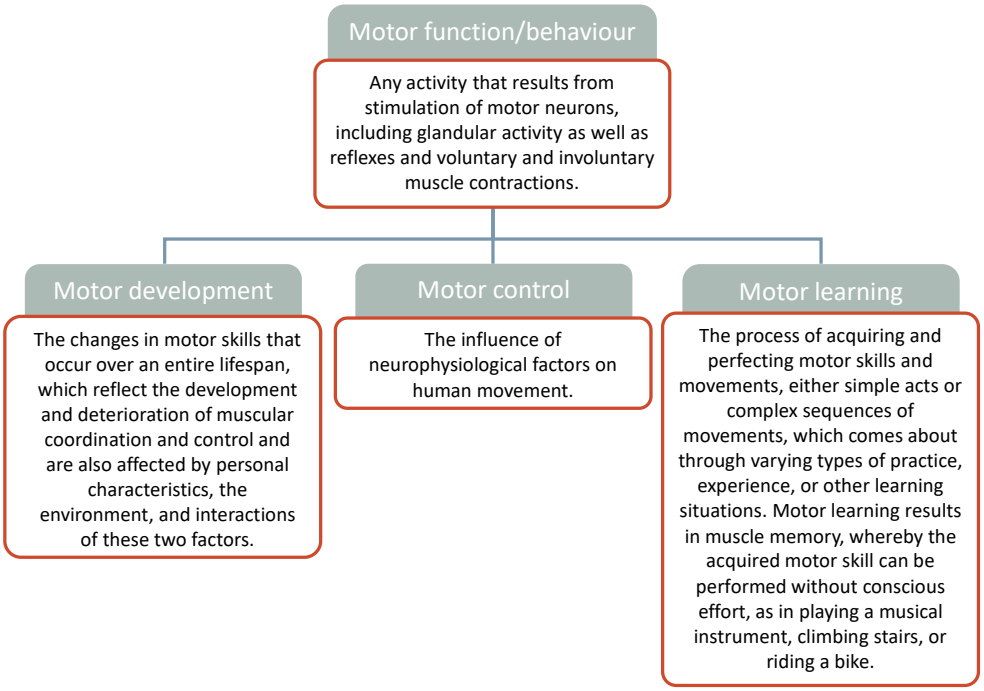


Fig. 1.1. Definitions of motor behaviour, motor development, motor control, and motor learning. Motor development, control and learning are sub-disciplines of motor behaviour, also known as motor function. (Definitions obtained from the American Psychological Association Dictionary [3]).

MOTOR CONTROL

Motor control addresses the underlying processes of movement [4], the organisation and control of functional movement [6]. Different motor control theories and models have been described (for an overview see [6]). One of these models, the so called conceptual closed-loop model (Figure 1.2), represents the motor control of skilled performance [9]. It emphasises the role of feedback on different levels. Assuming, a person wants to lift up a cup and drink from it, the stimulus for this is the person's thirst. The stimulus identification phase consists of the recognition of the sensory information, also from the environment [9] (e.g. recognition of thirst, visual input about the availability of a cup of water, etc.). Depending on the environmental situation, an appropriate response is selected [9] (e.g. the cup is positioned to perform the prior selected responses [9] (i.e. reaching to the cup, grabbing it and moving it to the lips, drinking, swallowing) and the physical prerequisites such as postural control are parameterised (stage of movement programming, Figure 1.2)

[9]. Consequentially, a motor program, which in the end will control the movements, is retrieved.

Motor programs are stored representations of movements and control centrally the various degrees of freedom that are involved when an action is performed [10]. Movements with the same invariant features (i.e. spatial and topographical characteristics [11], relative force and timing and sequences of the movement [12]) are considered to belong to the same motor program [12, 13]. The activated motor program sends the movement stimulus to the corresponding muscles via the spinal cord and peripheral motor nerves.

In the so-called closed-loop model of motor control [9], sensory feedback plays a pivotal role, as it presents the real world situation after the movement is executed. This real world situation then is compared with the anticipated feedback based on the initially programmed movement (Figure 1.2). The feed-forward information of the anticipated feedback represents the anticipated sensory consequences that should be received if the movement was carried out correctly [9]. As we are able to stop and correct a movement during the execution in case of an error, clearly, this comparison between planned movement and feedback must happen constantly throughout the whole movement. In the drinking-from-a-cup example, the feedback comparison would be about whether the hand has reached and grabbed the cup, the cup has reached the mouth and the water was swallowed the right way. In a perfect world, the whole movement is programmed perfectly, the movement is executed accordingly and the anticipated and the actual feedback both represent a successful drinking experience. Yet, the world we live in is not perfect. A joke has been told, the drinker might choke on the water because it is known that simultaneous drinking and laughing do not go together. Unplanned events disturb the whole process, leading to differences between the anticipated and the actual feedback. It is easy to imagine that when repeating an action several times, at least some of these stages of motor control will improve. According to Schmidt & Lee, practice leads to increased automaticity, speed and accuracy, in analysing the different feedback information (stimulus identification), improvements in response selection and parametrisation (movement programming), more effective motor programs, more accurate and precise feedback on different levels, and more precise indications of correctness (Figure 1.2) [9]. The improvement of motor control processes through practice is the essence of motor learning.

MOTOR LEARNING

Everyone who has ever learnt how to dance or play basketball, for example, knows that the acquisition of new motor skills requires extensive practice to lead to experience-dependent change in performance. This process is also known as motor learning [14]. In contrast to changes occurring due to motor development, which are based on maturation, changes due to motor learning require practice [4]. Motor learning enjoys great interest in the research fields of psychology, movement- and neuroscience and is pivotal not only for sports coaches, musicians or athletes, to name a few, but also for physiotherapists [14]. Motor learning is defined as "[…] a set of processes associated with practice or experience leading to relatively permanent gains in the capability for skilled performance." p. 178 [9]. This widely accepted definition implies that additionally to practice, several processes are involved in motor learning. Both the central and the peripheral nervous system including all sensorimotor functions and organs (e.g. muscles as executive system) are included. Furthermore, the gain is permanent, and thus, observable at least over a certain time period. The "capability for skilled performance" relates to the potential ability to attain an environmental goal, with a maximum of certainty and a minimum use of energy or time and energy [15]. To summarise this in a simplified manner, if the chosen environmental goal of a motor task can be reached with increased certainty, with less energy or time consumption due to practicing this very task, motor learning has occurred.

Different models have been established to describe motor learning from a superordinate point of view. One example is a systems approach, determining the person, the task, and the environment as systems influencing or enabling motor learning (Figure 1.3) [4]. Changes in one or more of these systems lead to different motor learning processes. Two different persons learning the same task in an identical environment might show different outcomes, due to their individual prerequisites, motivation or previously learnt skills.

The same person might need different settings (i.e. environments) to achieve the best possible outcome. When learning to dance, for example, a mirror positioned in front of the dancer provides direct visual feedback. When practicing with the back to the mirror, the dancer needs to rely on proprioceptive feedback. Different environments might lead to different motor learning outcomes in the same individual and enable the transfer of a learnt skill to another environmental setting or situation (e.g. learning a task in a laboratory situation or in a daily life environment). The task itself can be complex, simple, linked to an activity of daily life or be an artificial task. What is a motivational task for one person can be very boring for another person. The three systems environment, task, and individual are intermingled and cannot be considered separately from each other as a change in one of them, might requires a change in the others to enable motor learning.

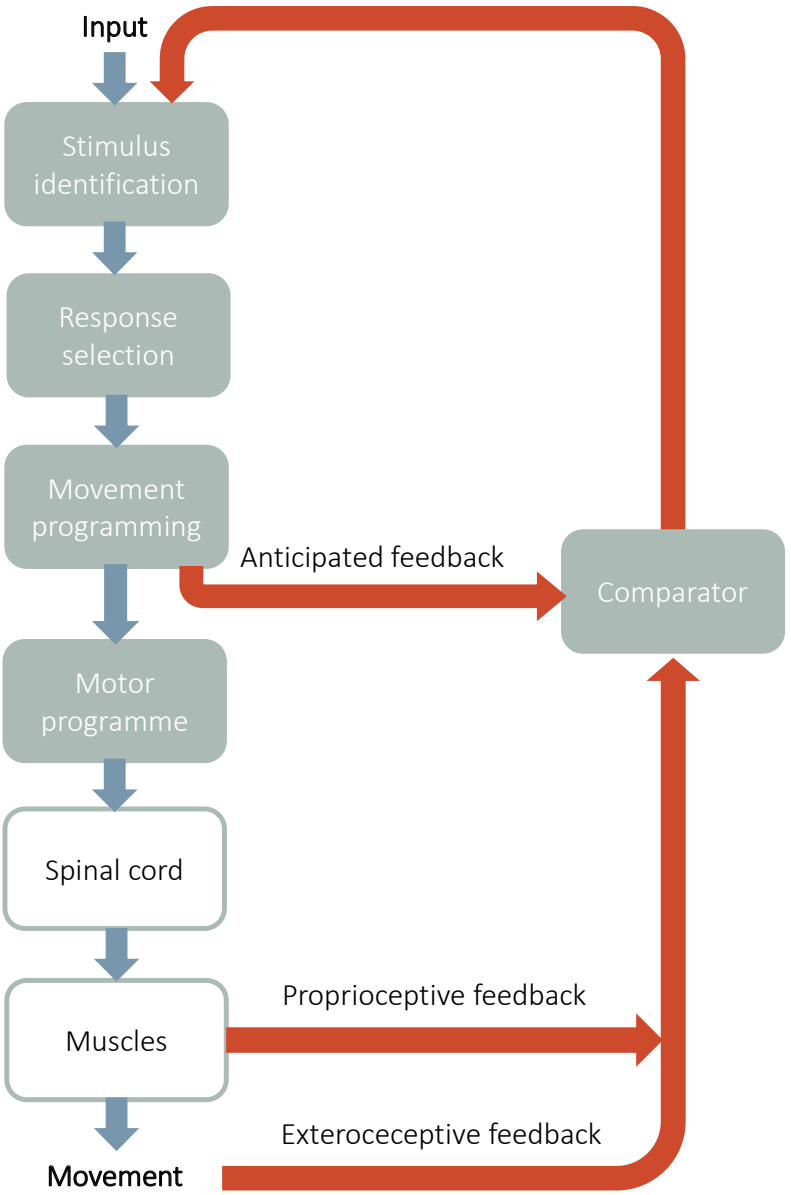


Fig. 1.2. Conceptual model of motor control. Different stages of motor control including the retrieval of a motor program and the sensory feedback channels, which feed the comparator with real-world information about the executed movement. Feed-forward information (anticipated feedback) based on the movement programming is compared to this actual feedback. All these components represent a closed-loop model of motor control. While improvement of motor control can happen on each level of the nervous system, the processes that change the most with practice are shaded in green.
Adapted from Schmidt & Lee, 2014 [9].

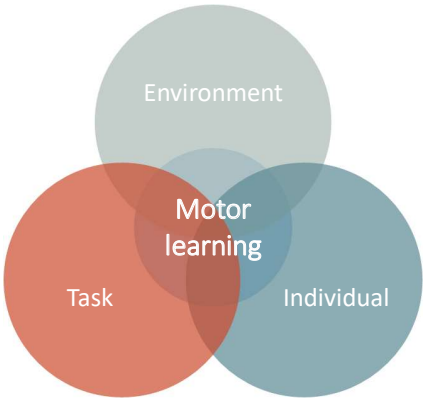


Fig. 1.3. Systems approach of motor learning. According to this approach, motor learning depends on the person, the task, and the environment. Adapted from Haibach-Beach et al., 2018 [4].

Motor learning is a matter of the brain

Learning in general requires neural plasticity, the ability of the brain to change. As Hebb`s learning rule, which was published first in 1949, proclaims, the efficiency of a synaptic transmission increases the more such a connection is used [16]. "What fires together wires together" is the famous quote of Hebb`s learning rule [16]. Hence, the synaptic strength increases with its use. Basically, this explains why practice is needed to induce any kind of learning. The increase of efficiency of the synaptic transmission is based on growth processes or metabolic changes in one or both synaptic cells [16]. Synaptic plasticity happens through an increase in the amount of neurotransmitters available and/or the increase of numbers of post-synaptic receptors for these neurotransmitters [17]. It has been observed also in motor learning, which seems to be based on a long-term potentiation-like process in the motor cortical synapses [18–21].

It is challenging to evaluate learning processes in the human brain, especially on the cellular basis as human brain cells are not so easily accessible [22]. Despite a large number of neuroimaging studies evaluating brain areas involved in motor learning, a consensus about consistently activated brain areas during motor learning is not reached [23]. On that account, a meta-analysis evaluating 70 motor learning neuroimaging experiments with adults was conducted [23]. It suggests that when learning a sensorimotor task besides the premotor cortex, the primary motor cortex and the supplementary motor area proper, the basalganglia and the cerebellum are predominantly active (Figure 1.4) [23]. According to the authors, the activation of the brain areas during learning of sensorimotor tasks is in line with the selective reinforcement of motor programs and the detection of sensorimotor prediction

error [23]. Motor programs are selected and organised during movement programming (see the conceptual model of motor control in Figure 1.2), which is most probably located in the basal ganglia [23, 24]. The detection of errors, the result of a comparison between the anticipated and the actual sensorimotor input, most certainly is located in the cerebellum [23, 25]. These imaging results seem to match the conceptual model of motor control and therewith the theoretical model of motor learning.

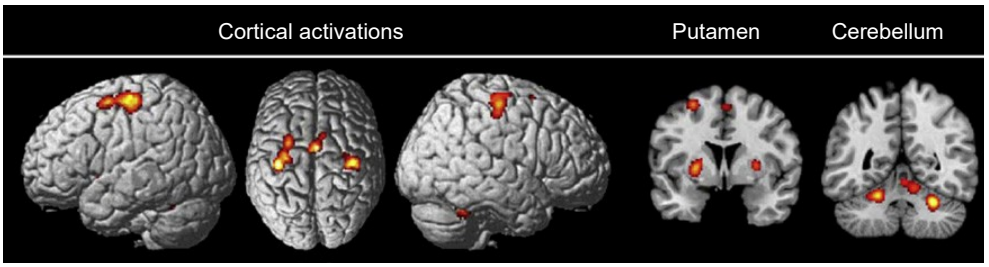


Fig. 1.4. Brain activation during learning of a sensorimotor task. Brain activation while practicing sensorimotor tasks: the left dorsal premotor cortex, bilateral primary motor cortices, supplementary motor area proper, bilateral putamen and bilateral/anterior medial cerebellum.
Adapted from Hardwick et al., 2013 [23].

Observing performance to infer whether motor learning occurs

The actual motor learning process is generally not directly observable since, as mentioned earlier, it involves alterations of the brain [9]. However, the products of motor learning can be observed, namely as a change in motor performance which is usually induced by motor learning [9]. Various parameters represent motor performance. Depending on the goal of the skill to be learnt, the parameter of interest could be the time needed to perform the task, the speed with which the task is performed or the number of errors made during execution. To present motor performance over time, it is common to relate the amount of practice and the parameter of interest in a so-called performance curve (Figure 1.5). A typical performance curve shows how the performance improves with the increasing number of practice trials. Usually, it starts with a steep improvement during the first trials, then the improvement becomes less, and finally stagnates in a plateau phase, reflecting the learning principle “law of practice” [26]. Dependent on the measure, this curve is increasing (e.g. increased speed or score) or decreasing (e.g. reduced time needed to perform task or a reduced number of errors) with improvement [9]. During the practice phase, the learner acquires a motor skill by practicing it. As it I shown in Figure 1.5, this phase is also known as acquisition. According to the definition mentioned earlier, motor learning requires a permanent change. Changes in performance occurring during acquisition are not necessarily permanent [4]. It is therefore important to consider the retention of a learnt skill, referring to the

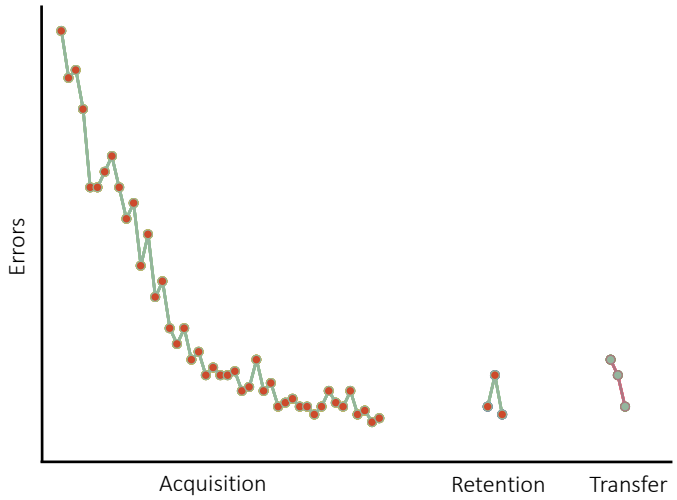


Fig. 1.5. Performance curve. With practice, the number of errors and the variability between the trials are reduced. In the initial phase of the acquisition this reduction is greatest. Later, to reach the same amount of improvement, more repetitions are needed and a plateau phase is reached. Orange dots and green lines represent the number of errors per trial of the practiced task and its retention, green dots and pink lines represent the errors per trial of the transfer task.

ability to reproduce the improved or newly learnt skill at a later time point, after the acquisition phase.

Additionally, to use a new skill in other environments, in other tasks or situations, the ability to transfer the skill to another situation is required. During all these features of the motor learning process, changes can be observed. Yet, it needs to be kept in mind that only a permanent long term change is referred to as motor learning [9]. Changes observed during practice should rather be referred to as momentarily improvement in motor performance.

All this knowledge about motor learning is not only important for highly skilled athletes and their coaches but also for people who need to learn and relearn how to move and their therapists. Motor learning is the basis of neurorehabilitation [27] and it is essential that the principles of motor learning are applied appropriately to therapies in order to achieve the best possible results. Yet, neurorehabilitation is not only a matter for adult patients with neuromotor impairments, but also for children. However, it is self-evident that findings from neurorehabilitation research in adults cannot completely be transferred to children [28]. Depending on the time point of neural maturation and development when the injury of the brain happens, the prerequisites for recovery are completely different [28]. Compared to adult people with brain impairments, the children's brains are more plastic [29]. What

sounds rather positive is not alone a guarantee for better recovery. Plasticity can be maladaptive and result in secondary disorders aggravating the situation like epilepsy [28]. Furthermore, several brain areas are especially vulnerable during development [30]. It has been stated that the benefit of the more plastic brain in children has to be weighted up against the unfinished development [28].

PAEDIATRIC NEUROREHABILITATION

While there are approximately 37 centres providing neurorehabilitation for adults in Switzerland [31], there is only one for children. In 2018, almost 200 young patients attended and finished an interdisciplinary inpatient rehabilitation programme at the Swiss Children's Rehab of the University Children's Hospital Zurich. The diagnoses varied but the majority of the children have a main diagnose of neurological origin (Figure 1.6).

An impairment of brain areas involved in motor control due to either congenital or acquired brain injuries results in several neuromotor disorders. Hypo- or hypertonia of the muscles, muscle weakness, sensory loss, involuntary movements or ataxia are just some of the symptoms. Cerebral palsy, for example, is associated with disorders of movement and posture related to defective movement coordination and/or muscle tone due to the damage of the immature brain [32, 33]. Depending on the severity of the impairment, the independence in everyday life is affected from completely dependent to a high level of independence.

Different approaches throughout different therapies (e.g. physiotherapy, occupational therapy, speech and language therapy, psychology, neuropsychology, sports therapy, or additional therapies like robotics for upper or lower extremities, aquatic therapy or hippotherapy) are followed in order to attain the goal of rehabilitation. Manual ability, for example, has a huge impact on independence in self-care [34]. When the main problem is based on an impairment of manual ability, the individual rehabilitation goal is set together with the child and the parents within this area. Depending on the type and severity of the impairment, treatment might also include the use of aids (for example wheelchairs, walking aids, aids for eating or bedding), the adaptation of splints or special shoes. Many of the children with brain impairments undergoing neurorehabilitation have affected upper limb functions. The prevalence of upper limb involvement for example in children with cerebral palsy is 83% [35]. As mentioned before, upper limb impairments have a huge impact on independence in everyday life. Therefore, it is of uttermost importance to pay attention to the different upper limb functions and their potential involvement. In addition to conventional therapies to improve upper limb functions, robot-assisted upper limb training is applied. It provides the potential of a highly repetitive training [36, 37], which is one great advantage as the increase of intensity and repetitions go together with an increased therapy impact in paediatric neurorehabilitation

[38]. Furthermore, the specificity of the task and the motivation of the patient to learn this task, are relevant aspects [9]. With the implementation of new technologies (e.g. robotic exoskeletons) in paediatric neurorehabilitation [39], a new way of incorporating these two aspects into therapy has thus been found.

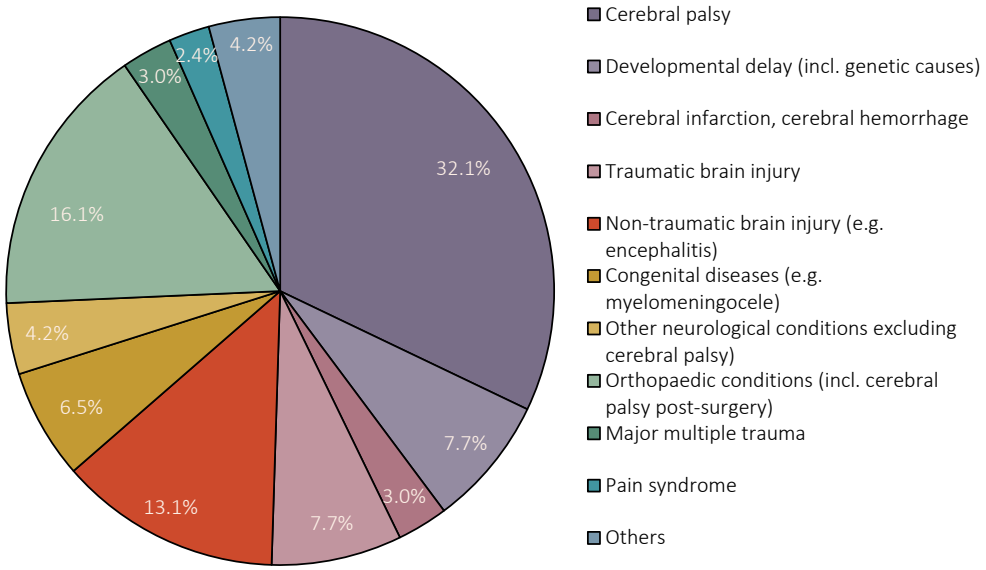


Fig. 1.6. Composition of the inpatients at the Swiss Children's Rehab, University Children's Hospital Zurich, in 2018. Nearly 200 patients finished rehabilitation in 2018. The majority of the patients has a neurological diagnosis.

Robot-assisted upper limb training in paediatric neurorehabilitation

Various rehabilitation technologies are used to train the different upper limb functions and in combination with exergames, a high number of repetitions is achievable in a motivational and playful way [40]. Usually, depending on the level of severity of the child's impairment, different levels of support are provided. Some tools even come with intelligent support modes which can be adjusted according to the patient`s actual performance (e.g. DIEGO®, Tyromotion, AT). Another advantage is, that these devices usually provide the assessment of different functions. [40]. An overview of some of the new technologies for the training and the assessment of upper limb movements used at the Swiss Children's Rehab are displayed in Figure 1.7.

The ChARMin robot

The ChARMin robot (Figure 1.8) is one of the robot-assisted therapy devices applied at the Swiss Children's Rehab in Affoltern am Albis for training upper limb function. ChARMin is an exoskeleton that has been developed and built at the Sensory-Motor Systems Lab (Federal Institute of Technology, Zurich, Switzerland) in collaboration

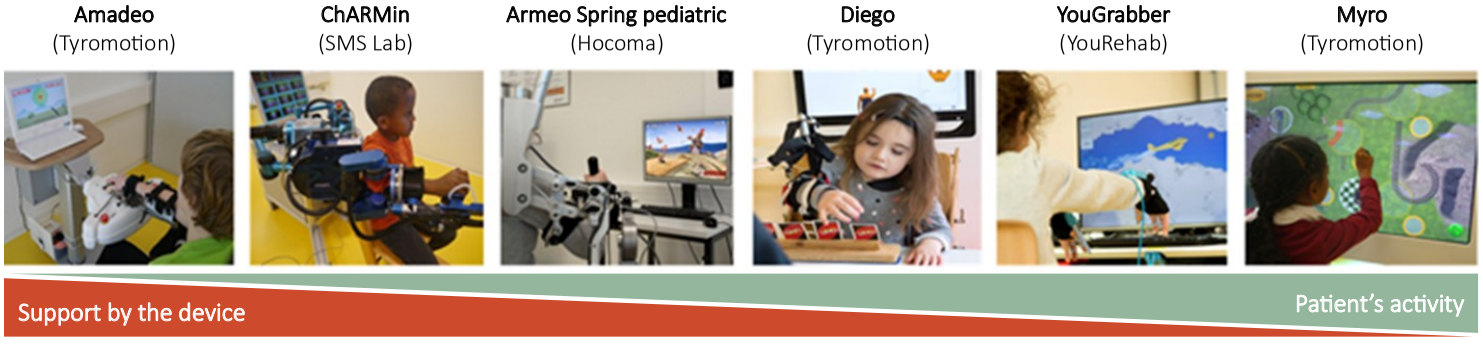


Fig. 1.7. New technologies to practice and assess upper limb function. The level of activity of the patient increasing from left to right while the level of support of the tool decreases from left to right.
F=function, A=activity, ICF = international classification of functioning disability and health.

with the Swiss Children's Rehab of the University Children's Hospital Zurich, Switzerland. ChARMin comes with two different distal modules, a small and a large one, and provides three support modes: fully supported, support as needed, and no support. Therefore, it can be used with children from five to 18 years, and different impairment levels [41]. The exoskeleton provides movements of shoulder horizontal abduction/adduction, flexion/extension, abduction/adduction, internal-/external rotation, elbow flexion/extension, forearm pronation/supination, wrist flexions/extension. Additionally, the hand module is sensitive to pressure and enables to include active hand closure in the training and assessments. Six exergames to train upper limb movements (e.g. goal-directed movements: playing tennis, whack-a-mole, shuffle board, capturing treasures and fish as a diver; movements to avoid obstacles: steering a plane or a space ship) are included. Preferences can be adjusted to single-joint or multiple-joint movements. A variety of parameters such as joint range, workspace, end effector range, mean speed or peak speed are recorded. This enables to evaluate the performance during training. Additionally, the therapist can teach the robot to perform a specific sequence of movements by moving the robot arm as desired. The robot then repeats the exact same sequence and performs a passive mobilisation of the child's arm to which it is attached. Furthermore, ChARMin provides six assessments to evaluate upper limb functions (active and passive range of motion, strength, resistance to passive movement, quality of goal directed movements (e.g. target precision, movement fluency), the dynamic tracking ability of the hand when following a circular movement on the screen, and the workspace, determining the maximally reached positions in six directions). These assessments enable to measure these functions objectively and potentially provide metric data to observe change through the course of rehabilitation. [36]. Regardless of whether a therapy is performed with or without robotic support, the active parts of many of these therapies are based on motor learning.

Fig. 1.8. The ChARMin. An exoskeleton to measure and improve upper limb functions in children and adolescents. ChARMin enables the training and assessment of upper limb functions and provides movements of the shoulder, the elbow, wrist and a pressure sensitive hand module enables to include fist closure as well. ChARMin comes with two differently sized distal modules and can be used with children of wide age and body size ranges.

Motor learning principles underlying paediatric neurorehabilitation

Yet, motor learning is not one complete package that can be added to therapy. Several motor learning principles have been described and need to be considered, leading to an individual recipe of learning. There is no generally exhaustive and valid list of motor learning principles, yet, usually they are divided between principles concerning the practice condition and the ones concerning feedback conditions [42–44]. Table 1.1 contains the most common motor learning principles. The need to repeat a motor skill to induce motor learning [9] is probably one of the least controversial principles.

Table 1.1. Motor learning principles

Domaine	Condition	Options
Practice	Amount	High vs. low number of repetitions/trials/sessions
	Distribution	Massed vs. distributed
	Variability	Constant vs. variable
	Schedule	Blocked vs. random order
	Attentional focus	Internal vs. external
	Task complexity	Simple vs. complex
Feedback	Type	Knowledge of result vs. knowledge of performance
	Frequency	High vs. low
	Timing	Immediate vs. delayed

Inexhaustible list of motor learning principles (adapted from Maas et al., 2008 [42]).

Repeat, repeat, repeat, …

Repetition seems to be the key to successful motor learning. Yet, how much repetition is sufficient to induce motor learning? Several studies from basic to clinical research have addressed this question. In rats, for example, 400 repetitions of a reaching movement per day during 10 days led to changes in the cortical movement representation (increase in the wrist and digit area, decrease in the elbow and shoulder area) [45]. Another study examined monkeys, who practiced approximately 600 repetitions of a pellet retrieval task per day during a minimum of 11 days [46]. The monkeys' cortical movement digit representations increased while their movement wrist/forearm representational zones decreased after the practice [45]. Adult participants after stroke improved their motor performance after 150 repetitions of both, a continuous tracking task and a serial reaction time task [47]. For a successful neurorehabilitation after stroke, it is recommended to provide highly intensive, repetitive task specific practice with feedback on performance [48]. However, in the clinical setting (occupational and physical therapy), adult patients with stroke practiced an average of 38.8 (± 30.9) repetitions of active upper limb exercise movements and 12.0 (± 12.3) purposeful upper limb exercise movements per session [49]. The same recommendation for the number of necessary repetitions probably also applies to children undergoing neurorehabilitation. Unfortunately, it is likely that also the number of repetitions in paediatric therapies is similar to the

ones observed in adult therapies. Hence, there is a large discrepancy between the recommendation and the actual implementation. So, what are the reasons for this?

The difficulty in achieving a high number of repetitions is the motivation or the lack thereof. Children develop their functions playfully [50] and play is a natural part of childhood [51, 52]. Therefore, play should be included in practice. Children with cerebral palsy practiced on average 3010 repetitions of a point-to-point movement during 70 minutes, divided in three sessions on three consecutive days, of upper limb exergame training with the paediatric Armeo® Spring (Hocoma AG, Volketswil, Switzerland) [36]. They showed successful acquisition, transfer, and retention of upper limb skills [36]. Hence, the issue of the large number of repetitions could be addressed by implementing robotic devices. While with robotic devices in general the implementation of a large amounts of repetitions could be facilitated, another important aspect should not be neglected: Children need to be interested, they need to have fun when practicing. To provide fun practice sessions and to keep the children engaged, a variable practice schedule is required. Therefore, during therapy sessions, usually several tasks or variations of one task are practiced. Variability (i.e. different tasks within the same session) is also suggested to be a facilitator of motor learning [53]. Yet, how should this variability be organised? The order in which these different skills are practiced, seems to be relevant [54]. Different orders of practicing several tasks have different impacts on the motor learning outcome [12].

Contextual interference

When several motor tasks or variations thereof are practiced together, the so called contextual interference effect occurs [12]. The different tasks functionally interfere with each other, leading to specific outcomes. The level of the interference depends on the practice order, an often referenced motor learning principle (Table 1.1 [42–44]). When referring to the contextual interference effect, generally, two practice orders are mentioned, the blocked practice order and random practice order. Blocked practice order means that one task is practiced several times before switching to the next task (A-A-A-A-B-B-B-B-B-C-C-C-C-C). Random practice means practicing the tasks in a random order (A-A-C-B-C-B-C-A-A-B-C-B). High contextual interference which occurs when practicing in random order, leads to better performance at retention and transfer but worse performance during acquisition compared to practicing under low contextual interference (blocked order) [55]. Figure 1.9 shows a schematic summary of the contextual interference effect.

This effect was firstly observed in the field of verbal learning and described by Battig in the 1960ies and 70ies [13, 56, 57]. Several years later, it was verified in the area of motor learning [55]. The contextual interference effect has mainly been assessed in fields of sports in healthy adults. There are some studies also assessing the contextual interference in children with and without disabilities, however, results are differing and study quality is generally low [58].

	Practice order	Performance		
		Acquisition	Retention	Transfer
Blocked	A-A-A-A-B-B-B-B-C-C-C-C	☺	☹	☹
Random	A-A-C-B-C-B-C-A-A-B-C-B	☹	☺	☺

Fig. 1.9. The contextual interference effect. If several motor tasks are practiced during the same session, blocked practice order leads to better motor performance during acquisition but to worse performance during retention and transfer compared to random practice order. A, B, and C representing three different motor tasks, ☺ refers to better performance, ☹ to worse performance, compared to the other practice order.

When practicing three variations of a visuomotor task over three days, healthy adults showed a decreased gamma-aminobutyric acid level in the occipital cortex as a result of random practice suggesting neuroplastic changes in this region [59]. There seems to be a difference on the adult brain's chemical level between blocked and random practice. It has been suggested that high contextual interference results in better learning because it requires more cognitive involvement compared to low contextual interference [60]. Accordingly, two principle hypotheses about the underlying processes have been established: the elaborative-processing hypothesis [61] and the forgetting-reconstruction hypothesis [62] (Figure 1.10). The elaborative-processing hypothesis is based on the idea that during random practice constant comparisons within and between the trials would lead to a deeper elaboration of the tasks compared to blocked practice [61]. This results in a more comprehensive and retrievable memory trace [60]. The forgetting-reconstruction hypothesis explains the benefit of random practice with forgetting of the before established action plan when performing a different task during the subsequent trial [60]. This leads to a strengthened memory consolidation compared to blocked practice [62].

A functional magnetic resonance imaging (fMRI) study, observed the brain activity during random or blocked practice of a bimanual visuomotor tracking task in adults [63]. Different movement directions and frequency ration (ratio of movement speed of the hand and the cursor) were applied as variations [63]. The authors hypothesized that random practice was more challenging and there would be a higher recruitment of the brain regions involved in bimanual task planning and execution that would be expected according to the reconstruction hypothesis [63]. Also, more involvement of sensory processing areas (i.e. mainly visual processing and integration regions, middle temporal region) for intratask-comparisons was anticipated during random practice, reflecting the elaboration hypothesis [63]. The contextual interference effect was

reconstructed, results supported both hypotheses, and the authors concluded that practice schedules had an impact on the differential modulation of brain regions [63]. Random practice generated an extra challenge inducing functional network plasticity resulting in positive behavioural effects [63].

A study with 60 healthy adult volunteers assessed the underlying hypotheses of contextual interference [64]. Participants practiced three different elbow extension-flexion movement patterns, either in blocked or random practice order [64]. In each practice order group, a third of the participants received monophasic transcranial magnetic stimulation (TMS) pulses applied during the inter-trial intervals TMS, a third received no TMS and a third sham-TMS [64]. TMS (eliciting isolated elbow-flexion and producing a motor-evoked potential in the biceps brachii) and sham-TMS were applied contralateral to the moving arm [64]. The aim of the TMS was to modulate the inter-trial elaboration processes to evaluate which of the two hypotheses provides the better explanation for the contextual interference effect [64]. If the elaboration hypothesis would be responsible for the contextual interference effect, the advantage of random practice would decrease since the elaborative processing would be disturbed by the TMS pulses while the performance under blocked practice would not be changed [64]. As proof for the forgetting-elaboration hypothesis, the authors suggest the approximation of performance under blocked practice to the performance under random practice because the inter-trial disturbance through TMS would lead to reconstruction processes also during blocked practice [64]. The results showed that the benefit of random practice deteriorated under TMS disturbances, hence, results support the elaboration hypothesis [64]. However, the results also partly supported the forgetting-reconstruction hypothesis because the TMS pulses during blocked practice led to a non-significant trend inducing benefit for motor learning [64].

Both hypotheses basically propose concepts of different levels of cognitive involvement when learning with different practice orders and are not mutually exclusive. Forgetting and reconstruction can occur also combined with a deeper elaboration process consequently combining both hypotheses.

Another, third hypothesis entitled retroactive inhibition hypothesis is based on a negative effect occurring during blocked practice rather than a benefit of the random practice order (Figure 1.10). It has been built on observations suggesting that the blocked order group performs worse at retention if the retention test task was the task which has been practiced at the beginning of the practice session compared to if the retention test task is the last practiced task [65]. Accordingly, the blocked order group would have a disadvantage, their previous memory would be inhibited due to the blocked practice order [66]. However, this hypothesis would only explain the difference between the practice orders in the performance during retention tasks and not during transfer and acquisition. Besides, in adults practicing goal-

oriented movements with a robotic manipulandum, which were disturbed through implemented force fields, the retroactive inhibition hypothesis has been rejected [66].

The physiological background of the contextual interference effect has also only been evaluated in adults. Disregarding its physiology, it would be interesting to know, whether the effect also works in children, especially in children with motor impairments due to brain lesions. It has been suggested in many research papers that children are not just small adults (e.g. [67]). Although this has often been stated in connection with medication, it applies to any kind of treatment and intervention.

Hypotheses	Elaboration hypothesis	Forgetting/reconstruction hypothesis	Retroactive inhibition hypothesis
Reasoning	Advantage of random order over blocked order practice Constant comparisons between and within the trials at random order leads to deeper elaboration	Advantage of random order over blocked order practice Forgetting of the before established action plan when practicing in random order leads to strengthened memory consolidation	Disadvanatge of blocked order practice Previous memory inhibited due to blocked practice
Results	Worse performance of the random order group during acquisition but better perfomance of the random order group during retention and transfer compared to the blocked order group	Worse performance of the random order group during acquisition but better perfomance of the random order group during retention and transfer compared to the blocked order group	Worse performance of the blocked group, compared to random order group, during retention if the retention task was the one practiced first during acquisition

Fig. 1.10. Hypotheses of the contextual interference effect. Three different hypotheses potentially underlying the contextual interference effects are displayed with their reasoning and the results of the effect that can be explained with the respective hypothesis.

PERSONAL MOTIVATION

All the topics of the preceding paragraphs had an influence on my PhD thesis. I was always very much interested in motor learning in general, be it as a physiotherapy student hearing of Bernstein's "repetition without repetition" for the first time, as a dancer being fascinated about factors that can influence the learning process, or more recently as a mother being astonished and overwhelmed by the progress of a child's motor development. This interest was also always self-evident in my work as a paediatric physiotherapist.

Rehabilitation therapy schedules are usually planned according to the timely and personnel resources. Quite early in my career, I was wondering whether it would be more efficient to provide therapy more topic related, meaning to schedule the different therapies in terms of a specific problem or a specific body area. For example, would a week focusing completely on upper extremity functions in all therapies be more successful compared to a week of rather arbitrary switching between different therapies with different topics and goals? Initially, this was also the question resulting in me doing a PhD thesis. However, it became clear to me, that the practice order even starts one step earlier, not between but within the therapy sessions: Does it make more sense to practice one task until proficiency is reached before starting with the next task or is practicing several tasks in a mixed or random order the right choice concerning the motor learning outcome? In my experience, it is very common to practice several different tasks within the same therapy session. Children are able to concentrate on one thing for quite some time but sooner or later, the task needs to change for the sake of motivation and interest. For all people involved, there is nothing as dissatisfying as an unmotivated and bored child in therapy. With my work at the Swiss Children's Rehab, I encountered robot-assisted therapy devices and gained quite some interest in the technical part of physiotherapy. Also, because advantages of quite easily measuring different functions related to movement were obvious, it was quite clear, that robotics will play a role in my PhD project.

I came across the more physiological parts of the topics of this introduction during the lectures within the PhD programme at the Centre for Neuroscience at the University of Zurich and the Federal Institute of Technology, Zurich. The background of motor learning has been assessed a lot in various fields in research and still, there is a great unknown, which always has attracted me. The following outline and the aims of my PhD thesis are a logical conclusion of my interests.

AIMS AND OUTLINE OF THIS PHD-THESIS

The general aim of this PhD-thesis is the investigation of the contextual interference effect in children with motor impairments due to brain lesions. Firstly, we conducted a systematic review to provide an overview over the existing evidence about the

contextual interference effect in children and in children with brain lesions (chapter 2). As evidence was lacking, the main aim became to design a randomised controlled trial evaluating the contextual interference effect in children with brain lesions. As robot-assisted devices in paediatric neurorehabilitation are advantageous and commonly applied at the Swiss Children's Rehab, we decided to perform the interventions with a robotic device. The ChARMin robot was chosen since it was thought to enable the training also with more severely affected participants, which seemed promising. Furthermore, it was a new device and the opportunity to work closely with the development engineer would facilitate technical adaptations or new developments if needed. Since ChARMin only recently has been developed and built, first, information about the psychometric properties of the ChARMin assessments was needed, as these assessments would serve to provide outcome parameters for the randomised controlled trial. Hence, secondly, a study evaluating relative and absolute reliability of the ChARMin assessments was conducted (chapter 3). Thirdly, some test trials with the existing exergames were conducted and showed that modified exergames would be needed to be able to assess the contextual interference effect with the ChARMin (chapter 4). According to the requirements, the ChARMin engineer programmed different versions of a new exergame. As no comparable study has been performed so far and our main objective was to evaluate the feasibility of a potential future randomised controlled trial, we decided to go for a pilot trial. Hence, and fourthly, we elaborated a detailed study protocol for a randomised controlled pilot study evaluating blocked versus random practice order of robot-assisted exergames training upper limb reaching movements in children with brain impairments (chapter 5). Finally, we conducted the pilot trial. We evaluated the feasibility of the randomised controlled trial with different predefined criteria (chapter 6). The overall discussion (chapter 7) deals with the main findings and methodological aspects of the studies, explains the different decisions in advancing during the whole project, and concludes with implications for future research.

REFERENCES

1. Young SN. The neurobiology of human social behaviour: An important but neglected topic. J Psychiatry Neurosci. 2008;33:391–2.
2. Bloom P. Does religion make you nice? Does atheism make you mean? SLATE - Human Interest. 2008; November 7.
3. American Psychologial Association Dictionary of Psychology. Motor function. https://dictionary.apa.org/motor-function.2020.
4. Haibach-Beach PS, Reid GD, Collier DH. Motor Learning and Development. 2nd editio. Champaign, IL: Human Kinetics; 2018.
5. American Psychologial Association Dictionary of Psychology. Motor behavior. 2020. https://dictionary.apa.org/motor-behavior.
6. Cech DJ, Martin S 'Tink'. Functional movement development across the life span. 3d edition. St. Louis, MO; 2012.
7. Largo RH. Babyjahre- das andere Erziehungsbuch. München: Piper Verlag GmbH; 2000.
8. Adolph KE, Franchak JM. The development of motor behavior. Wiley Interdiscip Rev Cogn Sci. 2017;8.
9. Schmidt RA, Lee TD. Motor learning and performance. From principles to application. 5th editio. Champaign; 2014.
10. Schmidt RA, Lee TD. Motor control and learning: a behavioral emphasis. 4th edition. Champaign: Human Kinetics; 2005.
11. Bernstein NA. The co-ordination and regulation of movements. Oxford: Pergamon Press; 1967.
12. Magill RA, Hall KG. A review of the contextual interference effect in motor skill acquisition. Hum Mov Sci. 1990;9:241–89.
13. Battig W. The flexibility of human memory. In: Cermak L, Craik F, editors. Levels of processing in human memory. Hillsdale; 1979. p. 23–44.
14. Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM. Motor learning.Compr Physiol. 2019;9:613–63.
15. Guthrie ER. The psychology of learning. Harper: Peter Smith Pub Inc; 1952.
16. Hebb DO. The organization of behavior: A neuropsychological theory. 2nd edition. Mahwaw, NJ: Lawrence Erlbaum Associates, Inc., Publishers; 2002.
17. Queensland Brain Institute. Long-term synaptic plasticity. The University of Queensland, Australia. 2017. https://qbi.uq.edu.au/brain-basics/brain/brain-

physiology/long-term-synaptic-plasticity.

18. Cantarero G, Lloyd A, Celnik P. Reversal of long-term potentiation-likeplasticity processes after motor learning disrupts skill retention. J Neurosci. 2013;33:12862–9.

19. Ziemann U, Iliac T V, Pauli C, Meintzschel F, Ruge D. Learning modifies subsequent induction of long-term potentiation-like and long-term depression-like plasticity in human motor cortex. J Neurosci. 2004;24:1666–72.

20. Stefan K, Wycislo M, Gentner R, Schramm A, Naumann M, Reiners K, et al. Temporary occlusion of associative motor cortical plasticity by prior dynamic motor training. Cereb Cortex. 2006;16.

21. Rosenkranz K, Kacar A, Rothwell JC. Differential modulation of motor cortcal plasticity and excitability in early and late phases of human motor learning. J Neurosci. 2007;27:12058–66.

22. Mansvelder HD, Verhoog MB, Goriounova NA. Synaptic plasticity in human cortical circuits: Cellular mechanisms of learning and memory in the human brain? Curr Opin Neurobiol. 2019;54:186–93.

23. Hardwick RM, Rottschy C, Miall RC, Eickhoff SB. A quantitative meta-analysis and review of motor learning in the human brain. Neuroimage. 2013;67:283–97.

24. Grillner S. Biological pattern generation: The cellular and computational logic of networks in motion. Neuron. 2006;52:751–66.

25. Rondi-Reig L, Paradis A-L, Lefort JM, Babayan BM, Tobin C. How the cerebellum may monitor sensory information for spatial representation. Front Syst Neurosci. 2014;8:1–13.

26. Snoddy GS. Learning and stability. J Appl Psychol. 1926;10:1–36.

27. Krakauer JW. Motor learning: Its relevance to stroke recovery and neurorehabilitation. Curr Opin Neurol. 2006;19:84–90.

28. Forsyth RJ. Back to the future: Rehabilitation of children after brain injury. Arch Dis Child. 2010;95:554–9.

29. Hensch TK, Bilimoria PM. Re-opening windows: Manipulating critical periods for brain development. Cerebrum. 2012; August.

30. Babikian T, Merkley T, Savage RC, Giza CC, Levin H. Chronic aspects of pediatric traumatic brain injury: Review of the Literature. J Neurotrauma. 2015;32:1849–60.

31. Emimo GmbH. StadtLandReha. 2018. https://stadtlandreha.ch.

32. Bax M, Goldstein M, Rosenbaum P, Leviton A, Paneth N, Dan B, et al. Proposed

definition and classification of cerebral palsy, April 2005. Dev Med Child Neurol. 2005;47:571–6.

33. Bax MCO. Terminology and classification of cerebral palsy. Dev Med Child Neurol. 1964;6:295–307.
34. Russo RN, Skuza PP, Sandelance M, Flett P. Upper limb impairments, process skills, and outcome in children with unilateral cerebral palsy. Dev Med Child Neurol. 2019;:1080–6.
35. Makki D, Duodu J, Nixon M. Prevalence and pattern of upper limb involvement in cerebral palsy. J Child Orthop. 2014;8:215–9.
36. Keller JW, van Hedel HJ. Weight-supported training of the upper extremity in children with cerebral palsy: A motor learning study. J Neuroeng Rehabil. 2017;14.
37. Duret C, Grosmaire A-G, Krebs HI. Robot-assisted therapy in upper extremity hemiparesis: Overview of an evidence-based approach. Front Neurol. 2019;10:1–8.
38. Bayón C, Raya R, Lerma Lara S, Ramírez Ó, Serrano J I, Rocon E. Robotic therapies for children with cerebral palsy: A systematic review. Transl Biomed. 2016;7:1–10.
39. van Hedel HJ, Aurich (-Schuler) T. Clinical application of rehabilitation technologies in children undergoing neurorehabilitation. In: Reinkensmeyer DJ, Dietz V, editors. Neurorehabilitation technology. 2nd editio. Switzerland: Springer International Publishing; 2016. p. 283–310.
40. van Hedel HJA, Lieber J, Ricklin S, Meyer-Heim A. Die praktische Anwendung von Exergames und virtueller Realität in der pädiatrischen Rehabilitation. neuroreha. 2017;09:35–40.
41. Keller U, Van Hedel HJA, Klamroth-Marganska V, Riener R. ChARMin: The first actuated exoskeleton robot for pediatric arm rehabilitation. IEEE/ASME Trans Mechatronics. 2016;21:2201–13.
42. Maas E, Robin DA, Hula SNA, Freedman SE, Wulf G, Ballard KJ, et al. Principles of motor learning in treatment of motor speech disorders. Am J Speech-Language Pathol. 2008;17:277–98.
43. Gill S V., Walsh MK. Use of motor learning principles to improve motor adaptation in adult obesity. Health (Irvine Calif). 2012;4:1428–33.
44. Bislick LP, Weir PC, Spencer K, Kendall D, Yorkston KM. Do principles of motor learning enhance retention and transfer of speech skills? A systematic review. Aphasiology. 2012;26:709–28.

45. Kleim JA, Barbay S, Nudo RJ. Functional Reorganization of the Rat Motor Cortex Following Motor Skill Learning. J Neurophysiol. 1998;80:3321–5.

46. Nudo RJ, Milliken GW, Jenkins WM, Merzenich MM. Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. J Neurosci. 1996;16:785–807.

47. Boyd L, Winstein CJ. Explicit information interferes with implicit motor learning of both continuous and discrete movement tasks after stroke. J Neurol Phys Ther. 2006;30:46–57.

48. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: A systematic review. Lancet Neurol. 2009;8:741–54.

49. Lang C, Macdonald J, Gnip C. Counting repetitions: An observational study of outpatient therapy for people with hemiparesis. J Neurol Phys Ther. 2007;31 March:3–10.

50. Whitebread D, Neale D, Jensen H, Liu C, Solis SL, Hopkins E, et al. The role of play in children' s development: A review of the evidence (research summary). The LEGO Foundation,DK. 2017; The LEGO Foundation,DK.

51. Kuschner D. Play is natural to childhood but school is not: The problem of integrating play into the curriculum. Int J Play. 2012;1:242–9.

52. Spodek B. What are the sources of early childhood curriculum? In: Spodek B, editor. Early childhood education. Englewood Cliffs, NJPrentice Hall.: Prentice Hall; 1973. p. 81–91.

53. Schmidt RA. A schema theory of discrete motor skill learning. Psychol Rev. 1975;82:225–60.

54. Hall KG, Magill RA. Variability of practice and contextual interference in motor skill learning. J Mot Behav. 1995;27:299–309.

55. Shea JB, Morgan RL. Contextual interference effects on the acquisition, retention, and transfer of a motor skill. J Exp Psychol Hum Learn Mem. 1979;5:179–87.

56. Battig WF. Facilitation and interference. In: Bilodeau EA, editor. Acquisition of skill. New York: Academic Press; 1966. p. 215–44.

57. Battig W. Intratask interference as a source of facilitation on transfer and retention. In: Voss J, editor. Topics in learning and performance. New York: Academic Press; 1972. p. 131–59.

58. Graser JV, Bastiaenen CH, van Hedel HJA. The role of the practice order: a systematic review about contextual interference in children. PLOS One. 2019;14.

59. Chalavi S, Pauwels L, Heise KF, Zivari Adab H, Maes C, Puts NAJ, et al. The

neurochemical basis of the contextual interference effect. Neurobiol Aging. 2018;66:85–96.

60. Lage GM, Ugrinowitsch H, Apolinário-Souza T, Mário Vieira M, Albuquerque MR, Novellino Benda R. Repetition and variation in motor practice: A review of neural correlates. Neurosci Behav Rev. 2015;57:132–41.

61. Shea JB, Zimny S. Context effects in memory and learning movement information. In: Magill, Richard A, editor. Memory and control of action. Amsterdam: North-Holland; 1983. p. 345–66.

62. Lee TD, Magill RA. The locus of contextual interference in motor-skill acquisition. J Exp Psychol Learn Mem Cogn. 1983;9:730–46.

63. Pauwels L, Chalavi S, Gooijers J, Maes C, Albouy G, Sunaert S, et al. Challenge to promote change: The neural basis of the contextual interference effect in young and older adults. J Neurosci. 2018;38:3333–45.

64. Lin C-HJ, Fisher BE, Winstein CJ, Wu AD, Gordon J. Contextual interference effect: elaborative processing or forgetting-reconstruction? A post hoc analysis of transcranial magnetic stimulation-induced effects on motor learning. J Mot Behav. 2008;40:578–86.

65. Shea JB, Titzer RC. The influence of reminder trials on contextual interference effects. J Mot Behav. 1993;25:264–74.

66. Thürer B, Gedemer S, Focke A, Stein T. Contextual interference effect is independent of retroactive inhibition but variable practice is not always beneficial. Front Hum Neurosci. 2019;13.

67. Klassen TP, Hartling L, Craig JC, Offringa M. Children are not just small adults: The urgent need for high-quality trial evidence in children. PLoS Med. 2008;5:1180–2.

CHAPTER 2

THE ROLE OF THE PRACTICE ORDER: A SYSTEMATIC REVIEW ABOUT CONTEXTUAL INTERFERENCE IN CHILDREN

Judith V Graser
Caroline HG Bastiaenen
Hubertus JA van Hedel

PLOS ONE 2019; 14(1)

ABSTRACT

Aim: We aimed to identify and evaluate the quality and evidence of the motor learning literature about intervention studies regarding the contextual interference (CI) effect (blocked vs. random practice order) in children with brain lesions and typically developing (TD) children.

Method: Eight databases (Cinahl, Cochrane, Embase, PubMed, Pedro, PsycINFO, Scopus and Web of Knowledge) were searched systematically with predefined search terms. Controlled studies examining the CI effect in children (with brain lesions or TD) were included. Evidence level, conduct quality, and risk of bias were evaluated by two authors independently. A best evidence synthesis was performed.

Results: Twenty-five papers evaluating TD children were included. One of these studies also assessed children with cerebral palsy. Evidence levels were I, II, or III. Conduct quality was low and the risk of bias high, due to methodological issues in the study designs or poor description thereof. Best evidence synthesis showed mainly no or conflicting evidence. Single tasks showed limited to moderate evidence supporting the CI effect in TD children.

Conclusion: There is a severe limitation of good-quality evidence about the CI effect in children who practice different tasks in one session, especially in children with brain lesions.

INTRODUCTION

Children with brain lesions, such as cerebral palsy (CP), frequently have to deal with impairments of the sensorimotor system, leading to restrictions in activities and independence which could affect participation in daily life [1]. Intensive therapeutic interventions are needed to address these limitations. Usually, several tasks or skills are practiced during single therapeutic sessions to cover a broad range of impairments and limitations and to keep the children engaged. However, learning one skill can be influenced by practicing another one during the same session [2]. This so-called contextual interference effect [3] has been established by Battig and has been described later in various motor learning studies. These studies showed that the contextual interference effect was low when different tasks are practiced in a blocked order, meaning that one task is practiced until it is learned before moving to the next [4]. A high contextual interference effect is achieved if different tasks are practiced in a random order [4]. Most evidence about the contextual interference effect has been obtained in healthy young adults with the intent of improving practice schedules in sports. In this population, a low contextual interference effect results in better acquisition but worse transfer and retention of task performance. The findings are the opposite if practicing with high contextual interference [4, 5].

For paediatric patients after rehabilitation discharge, it is important that learned tasks can be retained over time and generalised to other conditions or tasks. The evidence is lacking, though, whether this specific population would also benefit from a high contextual interference, and whether this can be achieved by practicing in a random order.

Several aspects seem to influence the contextual interference effect. Magill and Hall mentioned that task characteristics (e.g. non-laboratory tasks such as beanbag throwing vs. laboratory tasks such as coincident anticipation timing tasks) and subject characteristics like age or the level of experience are important, although it remains unclear how age exactly influences the contextual interference effect [5]. Therefore, it remains unclear what the optimal practice order in typically developing children (e.g. [6–8]) and in children with brain lesions undergoing neurorehabilitation is.

Some reviews about contextual interference exist [5, 9–12], but none of them included a systematic evaluation regarding the effects and quality of intervention studies in the field of paediatric motor learning, which limits the relevance for the field of paediatric neurorehabilitation. This shortage of knowledge is unfortunate since neurorehabilitation is based on motor learning principles [13], and therapeutic interventions could be improved by adhering to such principles [14]. As we assume that results from contextual interference studies involving typically developing children could be better generalised to children with brain lesions compared to results obtained from healthy adults, the objective of this systematic review was to investigate the evidence of contextual interference in children with congenital or

acquired brain lesions and typically developing children. The research question is the following: What is the evidence concerning the contextual interference effect for children with congenital and typically developing children?

METHODS

This review was conducted by following certain aspects of the guidelines provided by the American Academy of Cerebral Palsy and Developmental Medicine (AACPDM) [15] and supplemented by the risk of bias tool provided by the Cochrane Collaboration [16]. The procedure is described in detail below. Since no participants were required for this study, obtaining ethical approval was not necessary.

Inclusion and exclusion criteria

We defined inclusion criteria in line with PICO (**P**opulation, **I**ntervention, **C**ontrol, **O**utcome) and included studies assessing children (with congenital or acquired brain injuries and/or typically developing) in the age range between 1 and 18 years (Population). We included motor learning studies examining the contextual interference effect with a random practice order group (Intervention) and at least one blocked practice order group (Control). Any outcome evaluating the acquisition, retention, and/or transfer of the learned skill (Outcome) was considered selectable. JG defined the search terms based on PICO and HVH reviewed the search terms. The following search terms were used:

Population: 'child', 'children', 'childhood', 'paediatrics', 'adolescents', 'adolescence', 'youths', 'student', 'elementary', 'high school'.

Intervention: 'motor learning', 'skill learning', 'contextual interference', 'practice order';

Control group: 'blocked and random';

Outcome: 'performance', 'acquisition', 'retention', 'transfer', 'generalisation', and 'generalisability'.

Search terms were customised for each database including the use of MESH terms when applicable. We refrained from adding methodological criteria (e.g. randomisation procedures for group allocation) to get a broad overview of the existing literature. The search was performed by the first author on the databases Cinahl, Cochrane, Embase, PubMed, Pedro, PsycINFO, Scopus and Web of Knowledge (an example of a detailed search strategy is shown in S2.1 Table). The reference lists of original research papers and systematic reviews were screened for further eligible studies. The primary search was performed in March 2015 (for the period 1960 to March 2015) and updated in December 2016 (period 2015 to 2016).

We excluded studies that allowed a true practice phase (i.e. not a typical

familiarisation phase which normally consists of a few trials that are performed to have the participant give an idea about the skill to be learned) before the actual acquisition phase. There is a phenomenon called "learning to learn" [17] which describes the beneficial influence of prior practice experience on an unfamiliar motor task [18]. In humans, this phenomenon has been observed in visuomotor [18, 19], and cognitive tasks [20–22]. In a recent study with healthy young adults practicing a dynamic balance task, the "learning to learn" phenomenon could not be reproduced [23]. We included studies with a wide variety of motor tasks. Since there is no general accordance about the "learning to learn" phenomenon we decided to exclude studies with prior practice phase because this would affect the comparability with studies without such a practice phase. We also excluded conference papers, studies of which only the abstracts were available, unpublished dissertations, and studies in a language other than English or German.

Selection procedure

Firstly, JG and HVH read the titles and abstracts and decided upon eligibility independently. Secondly, the same authors read the full texts of the papers that were considered eligible and decided on final eligibility independently. In cases of disagreement, the authors discussed until consensus was reached.

Data extraction and analysis

JG summarised relevant data using a standardised data extraction sheet. Included were the type of study, participants (population, age, number per group), task, information regarding the acquisition, retention and transfer phases, including time points, duration, used outcome measures (e.g. anticipation timing task) and parameters (e.g. variable and random error), as well as the results. In case of incomplete reporting of patient characteristics or study procedure, we contacted the authors of the original publication.

We had planned to pool data when studies were comparable regarding populations, interventions, outcomes, and types of studies. If we were not able to follow this approach, due to heterogeneity of the studies, pooling within relevant subgroups was considered. When we would choose to refrain from pooling completely, because meaningful subgroups could not be built a best evidence synthesis would be performed using the levels of evidence described by Tulder et al. [24]. The results of each study would be rated as significant (favouring blocked or random order), inconsistent or not significant. Consistency of the results within one study would be given if 75% of the comparisons (e.g. measures, parameters, tasks) would provide similar results (e.g. random was better than blocked for the retention). Then the evidence of the different tasks (several studies per task, if possible) was rated according to the suggestions by Tulder et al. [24]: strong (consistent findings among multiple high quality randomised controlled trials (RCTs)), moderate (consistent

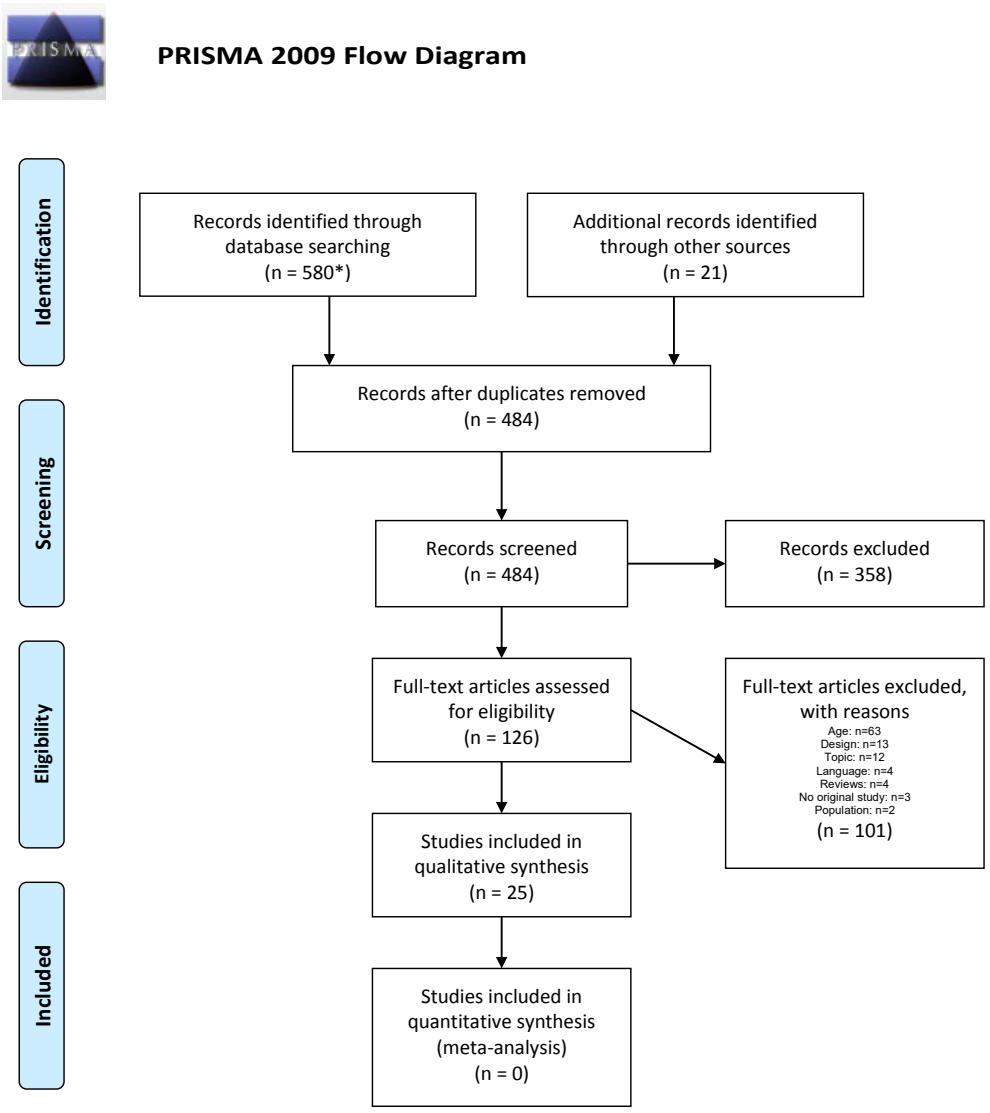
findings among multiple low quality RCTs and/or controlled clinical trials (CCTs) and/or high one high quality RCT), limited (one low quality RCT and/or CCT, conflicting (inconsistent fin-dings among multiple RCTs and/or CCTs; inconsistent findings among different parameters within one trial (if only one trial is available) or no evidence from trials (no RCTs or CCTs). Consistency of the studies assessing similar tasks would be given if more than 75% of the studies showed results in the same direction.

Methodological quality assessment

JG and CB assessed the level of evidence and the methodological quality of the eligible studies independently from each other, as recommended by the AACPDM [15]. The detailed descriptions of the evidence levels are displayed in S2.2 Table. The evaluation of the methodological quality included the seven aspects also described by the AACPDM [15] (for details see S2.3 Table).

We also evaluated the risk of bias. Bias is defined as any systematic error that results in an incorrect estimate of the true effect of an exposure on the outcome of interest. Bias can result in an over- or underestimation of the true value depending on the type of bias. We considered selection bias (i.e. sequence generation, allocation concealment), performance bias (i.e. blinding participants, personnel), attrition bias (i.e. incomplete outcome data), reporting bias (i.e. selective reporting), and other sources of bias (see also S2.4 Table). As bias is a potential threat to the trustworthiness of study results, the strength of a conclusion of a systematic review should be adjusted accordingly [16].

JG and CB rated the risk of bias according to recommendations described in the Cochrane Handbook for Systematic reviews of Interventions [16]. Discrepancies between the two authors were discussed until consensus was reached.



**includes records from primary search (for publications within time period 01.01.1960-31.03.2015) and search update (for publications within time period 2015-2016). There might be duplicates due to the overlap of the time periods (January to March 2015) which were removed during the second step.*

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

Fig. 2.1. Flowchart of the search process. Flowchart of the primary search (time period between 01.01.1960 and 31.03.2015) the updated search (time period of 2015 and 2016), and the inclusion and exclusion process.

RESULTS

Search results

The primary search in the databases led to 503 records (Fig 2.1). The 11 full texts that were excluded due to topic reasons were not motor learning studies or did not evaluate contextual interference. Thirteen full texts were excluded due to design issues (5 had no random order practice group, 4 had several practice orders within the same group (i.e. no parallel study design), 2 had a preparation phase, in which participants were allowed to practice for several sessions prior to the acquisition phase, 1 study had no blocked practice but a series of trials in blocks, and 1 had no blocked group but two random groups with different levels of variation).

By checking the references of eight reviews (four were found during the primary search, three within the references of these reviews and one by coincidence on google scholar) that focussed on contextual interference in general (not specifically for children), we could include two additional studies. The references of the original research studies contained no further eligible studies. The updated search resulted in one additional study (Fig 1).We included 25 papers in this systematic review. One paper presented three different experiments with three different samples [25]. This paper is handled as three separate studies in our review. Four of 27 studies assessed typically developing children as well as participants with disorders: Down's Syndrome [6], learning disabilities [26], mild mental handicaps [27], and CP [28]. Only the groups of typically developing children (i.e. with no diagnoses or disorders) were included for the best evidence synthesis. Concerning patient groups with congenital or acquired brain lesions, we included only the study in which children with CP learned to grasp unfamiliar objects [28]. Information of each study is presented in Table 2.1.

The methodological quality of the studies

Evidence levels (Table 2.2)

Most studies have an evidence level II or III, except for one, which had a level I [29]. Eight studies did not perform a randomisation [7, 14, 30–35], and were rated as level III. Two studies used cluster randomisation of school classes [36, 37]. One study randomly divided the participants into a complex and a simple task group and then further subdivided these groups into subgroups [14], but as this latter subdivision was not described, we did not consider it randomisation.

Quality of conduct (Table 2.2)

The methodological quality of the studies was low. Twenty-two studies received 0 out of 7 points. No study received a point for the questions 1 (allocation, randomisation), 2 (description and adherence of interventions), 5 (statistics), and 7 (appropriate methods the control confounding and bias). Two studies received 1 point (study by Jones & French [38], experiment 1 from Ste-Marie et al. [25]) and three studies received 2 points (study by Broadbent et al. [30], experiments 2 and 3 from Ste-

Marie et al. [25]).

Risk of bias

Most of the studies had a high risk of bias in all domains, except for Broadbent et al. [30] who defined the primary outcome measure and presented the results for all time points [30]. Therefore, we rated the risk of attrition bias as low.

Combining the study results by pooling the data in a meta-analysis was not appropriate since the studies were too heterogeneous considering the populations, types of motor tasks, intensities, time points (e.g. retention after five minutes, 24 hours or three weeks), and outcome measures. We also refrained from a subgroup analysis due to the low methodological quality and too small sample sizes of studies with sufficient relevant similarities.

Best evidence synthesis

The best evidence synthesis (Table 2.3) was conducted for the typically developing children. We grouped the studies according to the tasks they evaluated and received 15 task-specific groups. For most tasks, the evidence was conflicting or absent. Single tasks showed limited to moderate evidence supporting the contextual interference effect.

Acquisition: there was limited evidence for the benefit of blocked practice over random practice for dance step sequence [39], ball rolling, striking, and kicking [36], and a positioning motor task [40].

Retention: There was limited evidence for the benefit of random practice over blocked practice for throwing different balls [33] and playing tennis in a simulated environment [30].

Transfer: Moderately consistent evidence was found for the benefit of random practice over blocked practice for handwriting skills [25] and limited evidence for throwing different balls [33].

Table 2.1. Summary of the included studies

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Bertollo et al., 2010 [29]	Female high school students (mean age = 15.8 yrs., SD = 1.3 yrs.)	Dance step sequence (different sequences)	3 weeks 2 sessions / week (30 min each)	After 21 days	NA	Blocked > random ($p < 0.01$, $d = 0.90$, 95 % CI = 0.25- 1.56)	No significant group difference	NA
	Total: 40		Step sequence	Step sequence	NA			
	Blocked: 20 Random: 20		Score of spatial and temporal accuracy	Score of spatial and temporal accuracy	NA			
Bortoli et al., 1992 [30]	9th grade students (mean age = 14.6. yrs., SD = 0.7 yrs.)	Volleyball skills (bump, volley, surf)	6 weeks 1 session/week	After 1 week	After 1 week	Blocked = random ($F < 1.00$)	No significant group difference	Long transfer: random > blocked ($F_{3,48} = 2.97$, $p < 0.05$, all 4 groups) Short transfer: not reported
	Total: 52		Specified targets	Specified targets	Targets 1 meter (short transfer), 1 meter behind (long transfer)			
	Blocked: 13 Random: 13 Serial organisation: 13 Serial organisation (very high CI): 13		Scores of accuracy	Scores of accuracy	Scores of accuracy			
Broadbent et al., 2015 [31]	Intermediate - level tennis players (blocked practice group mean age = 12.9 yrs., SD = 1.6 yrs.; random practice group mean age = 13.2 yrs., SD = 1.6 yrs.)	Reaction to simulation: tennis skills (forehand groundstroke, forehand smash, forehand volley)	3 weeks 1 session / week (15 min each)	After 7 days	After 7 days	No main effect for group ($F_{1,16} = 0.10$, $p = 0.76$)	No main effect for group ($F^{1,16} = 2.9$, $p = 0.11$) Significant group x test interaction ($F^{1,16} = 6.03$, $p = 0.03$) random (mean ± SD = 71.7 ± 5.3 %) > blocked (mean ± SD = 63.3 ± 6.0 %)	Response accuracy: no significant group effect ($F_{1,14} = 0.03$, $p = 0.86$) Decision time: significant difference ($F_{1,14} = 7.19$, $p = 0.02$) random (mean ± SD = 98 ± 89 ms) > blocked (mean ± SD = 238 ± 118 ms)

(Continued)

Table 2.1. Continued

Study	Description			Acquisition		Retention		Transfer		Results		
	Participants		*Task/skill*	*Timeframe*		*Timeframe*		*Timeframe*				
	n total			*Ourtcome measures*		*Ourtcome measures*		*Ourtcome measures*		*Acquisition*	*Retention*	*Transfer*
	n group			*Outcome parameters*		*Outcome parameters*		*Outcome parameters*				
Broadbent et al., 2015 [31] *(Continued)*	Total: 18			Tasks themselves		Tasks themselves		Tasks themselves				
	Blocked: 9 Random: 9			Response accuracy = primary outcome		Response accuracy = primary outcome		Response accuracy Decision time				
Del Rey et al., 1983 [32]	Children, half of the sample experienced in sports requiring coincident anticipation (median age = 8.33 yrs., range = 6.50 yrs. to 10.67 yrs.		Anticipation timing task	1 session 60 trials		NA		Immediately after		Significant group difference blocked > random (probably for both, absolute and variable error, since "analyses for the two errors, for the most part, agreed" p. 582)	NA	Absolute error: no significant group difference ($p = 0.05$) (blocked: mean ± SD = 70 ± 45 ms; random: mean ± SD = 106 ± 62 ms) Variable error: no significant group difference ($p = 0.05$) (blocked: mean ± SD = 74 ± 25 ms; random: mean ± SD = 98 ± 42 ms)
	Total: 80			Task itself		NA		Task itself, new speed				
	Blocked: 40 Random: 40 20 girls and boys in each practice group			Absolute error Variable error		NA		Absolute error Variable error				
Duff et al., 2003 [28]	Children with hemiplegic CP (mean age = 10 yrs., SD = 1.8 yrs.) TD children (mean age = 10.4 yrs., SD = 1.7 yrs.)		Lifting unfamiliar objects	1 session 81 trials		Immediate: after 5 min Delayed: after 24 hours		NA		All participants: Grip force rate: object differentiation in the 9th block: blocked > random ($p < 0.05$) other blocks: not reported Load force rate: object differentiation in the 9th block: blocked > random ($p < 0.001$) other blocks: no significant group difference	All participants: no significant group difference CP: not reported TD: not reported	NA
	Total: 36			Tasks themselves		Tasks themselves		NA				

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Duff et al., 2003 [28] *(Continued)*	Children with CP: Blocked: 9 Random: 9 TD children: Blocked: 9 Random: 9		Load force rate Acceleration after lift-off Grip force Grip force rate	Load force rate Acceleration after lift-off Grip force Grip force rate	NA	*(Continued)* Acceleration after lift-off: object differentiation in the 1st block: blocked > random ($p < 0.01$) Acceleration between objects more similar for blocked vs. random group ($p > 0.05$) CP: not reported TD: not reported		
Edwards et al., 1986 [6]	Children with DS (mean CA = 18.1 yrs., SD = 2.5 yrs.; mean MA = 4.7 yrs., SD = 1.4 yrs.) Children without DS (mean CA = 5.8 yrs., SD = 2.0 yrs.)	Coincident anticipation timing task	1 session 4 x 16 trials	NA	After 10 minutes	Both groups (DS and no DS): Absolute constant error: no significant group difference Variable error: no significant group difference	NA	Inside transfer: Absolute constant error: no significant difference Variable error: significant group x training x block interaction ($F_{1,36}$=4.1) random > blocked (no DS) Outside transfer (both groups (DS and no DS): Absolute constant error: no significant difference Variable error: no significant difference
	Total: 40		Task itself	NA	Same task, different speeds (within transfer: speed within the trained speed range; outside transfer: speed outside the trained speed range)			
	Children with DS: 20 Children without DS: 20 n/practice group unknown		Absolute constant error Variable error	NA	Absolute constant error Variable error			

(Continued)

Table 2.1. Continued

Study	Description			Acquisition		Retention	Transfer		Results	
	Participants		*Task/skill*	*Timeframe*		*Timeframe*	*Timeframe*		*Acquisition*	*Retention*	*Transfer*
	n total	*n group*		*Ourtcome measures*	*Outcome parameters*	*Ourtcome measures*	*Ourtcome measures*	*Outcome parameters*			
Fialho et al., 2006 [33]	Skilled volleyball players (mean age = 16.3 yrs., SD = 0.67 yrs.)		Volleyball skills (tennis serve, float serve)	4 days 1 session / day (46 trials each)		NA	After 10 minutes After 24 hours (retention of the transfer)		Mean score: NA (between group results not reported) SD of the score: no significant group difference	NA	Mean score: Transfer: NA (between group results not reported) Retention of the transfer: random > blocked $H_{(0,1)} = 3.6$, $P < 0.05$) for the first block of trials SD of the score: Transfer: no significant group difference Retention of the transfer: no significant group difference
	Total: 10			Tasks themselves		NA	Asian serve (transfer and retention of the transfer)				
		Blocked: 5 Random: 5		Accuracy scores of the serves (means and SDs)		NA	Accuracy scores of the serves (means and SDs)				
French et al., 1990 [34]	9th grade students, enrolled in physical education (mean age/SD = NA)		Volleyball skills (forearm pass, overhead set, serve)	6 days 1 session / day (30 trials each		After 2 days	NA		No significant main or interaction effect	No significant main effect	NA
	Total: 139			Tasks themselves		Tasks themselves	NA				
		Blocked: NA Random: NA Random-blocked: NA		Scores for accuracy and force		Scores for accuracy and force	NA				
Gophna et al., 2007 [26]	Children with LD (mean age = 7.1 yrs., SD = 0.25 yrs.) Children without LD (mean age = 7.9 yrs., SD = 0.31 yrs.)		Mirror tracing task (different shapes)	1 session 36 trials		After 1 week	After 1 week		n errors: random > blocked (all participants, p=0.02) Error time: no significant group difference Total time: no significant group difference	n errors: no significant group difference Error time: no significant group difference Total time: no significant group difference	n errors: no significant group difference Error time: no significant group difference Total time: no significant group difference

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Gophna et al., 2007 [26] *(Continued)*	Total: 48		Tasks themselves (3 shapes)	Tasks themselves (3 shapes)	Same task, 2 different shapes			
	Children with LD: 24 Blocked: 12* Random: 12* Children without LD: 24 Blocked: 12* Random: 12*		n errors Error time Total time (to complete the task)	n errors Error time Total time (to complete the task)	n errors Error time Total time (to complete the task)			
Granda Vera & Montilla, 2003 [35]	6-yr.-old children (considered normal and healthy)	Throwing different balls	6 weeks 3 sessions / week (50 minutes each)	After 2 weeks	After 2 weeks	Vertical target: Tennis ball (3 m): random > blocked ($F = _{7.54}$, $p = 0.008$) Feather fly ball (3 m): random > blocked ($F = _{5.90}$, $p = 0.02$) Tennis ball (5 m): no significant group difference Feather fly ball (5 m): no significant group difference Horizontal target: Tennis ball (3 m): no significant group difference Feather fly ball (3 m): no significant group difference Tennis ball (5 m): no significant group difference Feather fly ball (5 m): no significant group difference Total score vertical target: random > blocked ($F = _{5.68}$, $p = 0.02$) Total score horizontal target: no significant group difference Total score all conditions: random > blocked ($F = _{3.68}$, $p = 0.05$)	Random (mean ± SD = 13.0 ± 4.47) > blocked (mean ± SD = 9.29 ± 3.62) ($F = _{16.26}$, $p = 0.001$)	Random (mean ± SD = 10.08 ± 54.01) > blocked (mean ± SD = 7.58 ± 2.72) ($F = _{11.79}$, $p = 0.001$)
	Total: 71		Tasks themselves	Tasks themselves	Same task, different ball, different distance			
	Blocked: 34 Random: 37		Performance score	Performance score	Performance score			

(Continued)

Table 2.1. Continued

Study	Description			Acquisition	Retention	Transfer	Results		
	Participants		*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total			*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group			*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Green et al., 1995 [36]	Female students (first year secondary school eastern England), right handed novices in racket sports		Hitting different balls with different rackets (tennis, squash, badminton, short tennis)	4 days (48 trials each)	After 1 minute, after 1, 4 and 8 days	After 1 minute, after 1, 4 and 8 days	No significant group difference	NA	Both dimensions out of range: (racket and target): random > blocked ($F = 6.84$, $p < 0.05$) One dimension out of range: (racket or target): blocked > random ($F = 10.07$, $p < 0.01$)
	Total: 48			Tasks themselves	Tasks themselves	Same task, different racket and / or different target			
	Control: 12* Specific: 12* Blocked: 12* Random: 12*			Error scores (target accuracy)	Error scores (target accuracy)	Error scores (target accuracy)			
Jarus & Goreover, 1999 [37]	Healthy children, no motor or cognitive deficits (mean age = 8.15 yrs., SD = 2.44 yrs.)		Throwing beanbags of different sizes to targets of different distances	1 session 30 trials	After 30 minutes	After 30 minutes	All participants: blocked (mean ± SD = 21.30 ± 9.66 cm) > random (mean ± SD = 26.30 ± 12.20 cm) $F_{1,81} = 5.29$, $p < 0.01$ 5-yr.-old: no significant group difference 7-yr.-old: blocked (mean ± SD = 22.11 ± 8.65 cm) > random (mean ± SD = 33.11 ± 11.17 cm) 11-yr.-old: no significant group difference	All participants: blocked (mean ± SD = 19.75 ± 10.29 cm) > random (mean ± SD = 25.30 ± 11.37 cm) $F_{3,108} = 3.73$, $p < 0.05$ 5-yr.-old: no significant group difference 7-yr.-old: blocked (mean ± SD = 17.17 ± 8.05 cm) > random (mean ± SD = 28.25 ± 8.34 cm) 11-yr.-old: no significant group difference	No significant group difference for neither all participants nor any of the age groups
	Total: 120			Tasks themselves	Tasks themselves	Same task, different target			

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Jarus & Goreover, 1999 [37] *(Continued)*	5-yr.-old: 40 Blocked: 10 Random 10 Combined:10 Control: 10 7-yr.-old: 40 Blocked: 10 Random 10 Combined:10 Control: 10 11-yr.-old: 40 Blocked: 10 Random 10 Combined:10 Control: 10		Distance from target	Distance from target	Distance from target			
Jarus & Gutman, 2001 [14]	Children from public school, no cognitive and motor deficits (mean age = 8.52 yrs., SD = 0.61 yrs.)	Throwing beanbags of different sizes to targets of different distances Simple task: different bag weights Complex task: different bag weights, and sizes, different target order	1 session 30 trials	After one day	After one day	Simple task: Total time: no significant group difference Accuracy score: not reported due to lack of significant results Complex task: Total time: blocked (mean ± SD = 10.83 ± 3.1) > random (mean ± SD = 14.94 ± 4.34) Accuracy score: not reported due to lack of significant results	Simple task: Total time: no significant group difference Accuracy score: not reported due to lack of significant results Complex task: Total time: no significant group difference Accuracy score: not reported due to lack of significant results	Simple task: Total time: no significant group difference Accuracy score: not reported due to lack of significant results Complex task: Total time: no significant group difference Accuracy score: not reported due to lack of significant results
	Total: 96		Tasks themselves	Tasks themselves	Same task, different bag (simple transfer), different bag and different target order (complex transfer)			
	Simple task: 48 Blocked: 16 Random: 16 Combined: 16 Complex task: 48 Blocked: 16 Random: 16 Combined: 16		Total time to complete each trial Accuracy score (not reported due to lack of significant results)	Total time to complete each trial Accuracy score (not reported due to lack of significant results)	Total time to complete each trial Accuracy score (not reported due to lack of significant results)			

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Jones & French, 2007 [38]	9th grade students from 3 high-schools, physical educational classes	Volleyball skills (underhand serve, forearm pass, overhead set)	9 days 30 trials / day (approximately 30 min)	After 2 days	NA	No significant group difference ($p>0.05$)	No significant group difference ($p>0.05$)	NA
	Total: 68 51 completed the whole procedure		Tasks themselves	Tasks themselves	NA			
	Blocked: 18* Random: 17* Blocked-random: 16*		Scores (accuracy of the volleyball skills)	Scores (accuracy of the volleyball skills)	NA			
Meira & Tani, 2003 [39]	Female students, secondary school, right-handed, volleyball Novices (mean age = 12.7 yrs., SD = NA	Volleyball skills (underhand serve, overhand serve, Asian floater)	8 sessions (2 / week) 36 trials / session	NA	Immediately after: Transfer 1 3 sessions (2 / week) 28 trials / session 1 week after: Transfer 2 1 session 12 trials	Precision scores on target: no significant group difference movement pattern quality scores: No significant group difference	NA	No significant group difference in neither transfer 1 nor 2, in neither of the parameters
	Total: 36		Tasks themselves, with knowledge of result	NA	Tasks themselves, without knowledge of result			
	Blocked: 18 Random: 18		Precision scores on target Movement pattern quality scores	NA	Precision scores on target movement pattern quality scores			

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Painter et al., 1994 [27]	Students enrolled in IDD classrooms at public junior high-school (mean age = 13.9 yrs., SD = NA) Gender- and CA-matched controls from same school district (mean age = 13.11 yrs., SD = NA)	Throwing beanbags with different throws (underhand throw, over-hand throw, hook throw	1 session 45 trials (3 x 15)	After 10 minutes	NA	Absolute error: IDD: no significant group difference Controls: no significant group difference Combined: significant trial block x schedule interaction ($F_{14,616}$ = 3.84, p < 0.001), significant error reduction in random group but not in blocked group Variable error: IDD: not reported Controls: not reported Combined: not reported	Absolute error: IDD: blocked > random (p < 0.05 for all comparisons) Controls: no significant group difference Combined: random > blocked ($F_{1,44}$ = 15.77, p < 0.001) Variable error: IDD: not reported Controls: not reported Combined: consistency random > blocked ($F_{1,44}$ = 7.94, p < 0.007)	NA
	Total: 48		Tasks them-selves	Tasks them-selves	NA			
	Children with IDD: 24 Blocked: NA Random: NA Control group: 24 Blocked: NA Random: NA *Equal number of males/females in each practice group*		Absolute error Variable error	Absolute error Variable error	NA			
Perez et al., 2005 [40]	Children (mean age = 10.5 yrs., SD = 0.6 yrs.)	Positioning motor task	1 session 60 trials	NA	Immediate transfer: 10 trials, after 5 minutes Delayed transfer: 20 trials, after 24 hours	Blocked > random (p < 0.01) in all acquisition trials except block 9 (no significant group difference)	NA	Immediate transfer: not reported Delayed transfer: random > blocked (p < 0.0001) for the first block

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Perez et al., 2005 [40] *(Continued)*	Total: 57	Positioning motor task	Task itself	NA	Task itself, different position, no knowledge of result			
	Blocked: 29 Random: 28		Absolute error	NA	Absolute error			
Pigott & Shapiro, 1984 [7]	Students from the Montessori school in Santa Monica, CA (age range = 6 yrs. 9 months to 8 yrs. 3 months)	Throwing bean bags (different weights)	1 session 24 trials	NA	Immediately after 3 trials	No significant group differences	NA	No significant group differences
	Total: 64		Task itself	NA	Task itself, different weights			
	Blocked: 16 Random: 16 Random-blocked: 16 Constant: 16 ***Groups balanced for sex and average age***		Absolute error	NA	Absolute error			
Pollock & Lee, 1997 [41]	Children, 7-yr.-old (mean age = 7.2 yrs., SD = 0.4 yrs.) Adults (mean age = 24.3 yrs., SD = 3.1 yrs.)	Propelling a wooden disk with the middle finger (adaptation of the Crokinole game)	1 session 90 trials (3 x 30)	Immediately after (after transfer test) 15 trials (3 x 5)	Immediately after 2 tests, 10 trials each	All participants: blocked (mean ± SD = 1.4 ± 0.8) > random (mean ± SD = 1.2 ± 0.8) significant main effect for order ($F_{1,44}$ = 5.16, p < 0.05) Children: blocked (mean ± SD = 0.8 ± 0.6) = random (mean±SD=0.8±0.7)	All participants: blocked (mean ± SD = 1.2 ± 0.7) < random (mean ± SD = 1.5 ± 0.9) significant main effect for order ($F_{1,44}$ = 4.13, p < 0.05) Children: not reported Adults: not reported	All participants: blocked (mean ± SD = 1.3 ± 0.8) = random (mean ± SD = 1.7 ± 0.9) $F_{1,44}$ = 3.76, p = 0.06 Children: not reported Adults: not reported

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Pollock & Lee, 1997 [41] *(Continued)*	Total: 48		Task itself	Task itself	Task itself, new starting point, new "bumper	*(Continued)* Adults: blocked (mean ± SD = 2.1 ± 1.0) > random (mean ± SD = 1.6 ± 1.0)		
	Children: 24 Blocked: NA Random: NA Adults: 24 Blocked: NA Random: NA *Group membership balanced by sex*		Scores (accuracy of the target)	Scores (accuracy of the target)	Scores (accuracy of the target)			
Saemi et al., 2012 [42]	Male elementary school students, low skilled in throwing tasks (mean age = 10.47 yrs., SD = 0.77 yrs.)	Throwing tennis balls from different starting positions to different targets	1 session 81 trials (3 x 27)	After 1 day (12 trials)	NA	Significant main effect for practice condition ($F_{2,33}$ = 4.19, $p < 0.024$, $\eta^2 = 0.203$) but no post-hoc pairwise comparison reported No significant practice condition x trial block interaction ($F_{16,164}$ = 0.44, p = 0.97)	No significant group difference	NA
	Total: 36		Task itself	Task itself	NA			
	Blocked: NA Random: NA Increasing: NA		Scores (accuracy of the target)	Scores (accuracy of the target)	NA			
Stambaugh, 2011 [43]	Beginning clarinet players (from 16 elementary schools, 5th or 6th grade, in five school districts from the northwest United States (mean age = NA, SD = NA)	Playing clarinet	3 sessions 18 trials / session	After 24 hours 3 trials	After 24 hours (after retention) 3 trials of each task	Speed: blocked > random (blocks 1, 2, 3) random > blocked (block 4 and further) ($F_{5,34}$ = 5.052, p = 0.001) Accuracy: no significant group difference (p = 0.28) Temporal evenness: no significant group difference (p = 0.35)	Speed: random > blocked ($F_{1,38}$ = 24.953, $p < 0.001$) Accuracy: no significant group difference (p = 0.44) Temporal evenness: no significant group difference (p = 0.06)	Speed: no significant group difference (p = 0.78) Accuracy: no significant group difference (p = 0.36) Temporal evenness: no significant group difference (p = 0.88)

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Stambaugh, 2011 [43] *(Continued)*	Total: 41		Task itself	Task itself	Task itself, other note sequence			
	Blocked: 22 Random: 19		Scores (speed, accuracy, temporal evenness)	Scores (speed, accuracy, temporal evenness)	Scores (speed, accuracy, temporal evenness)			
Ste-Marie et al., 2004 [25] Experiment 1	Children from 3 1st grade classes from 2 Catholic elementary schools (21 females: mean age = 6.3 yrs., SD = 0.43 yrs.; 23 males: mean age = 6.2yrs., SD = 0.55 yrs.)	Handwriting tasks	1 session (35 minutes) 72 trials (24 trials / symbol)	After 30 minutes	NA	No significant group difference blocked (mean ± SD = 2.42 ± 0.95) = random (mean ± SD = 2.26 ± 0.86) ($F_{1,42}$ = 3.2, MSE = 1.9, p = 0.08)	Random > blocked (for both retention groups) ($F_{1,42}$ = 4.2, MSE = 129, p = not reported)	NA
	Total: 44		Tasks themselves	Tasks themselves (either in blocked or random order)	NA			
	Blocked: NA Random: NA		Scores (scoring of the handwritten symbols)	Scores (scoring of the handwritten symbols)	NA			
Ste-Marie et al., 2004 [25] Experiment 2	Children, recruited from 2 schools (mean age = 6.90 yrs., SD = 0.51 yrs.)	Handwriting tasks	1 session (35 minutes) 72 trials	After 20 minutes After 24 hours	After 20 minutes After 24 hours (after the retention test)	Random > blocked (trial sets 2, 3, 4) Blocked = random (trial sets 1, 5, 6, 7, 8) ($F_{7,322}$ = 3.2, MSE = 0.67, p = not reported)	20 minutes retention: significant letter x group interaction ($F_{2,92}$ = 3.4, MSE = 0.090) Blocked: significant writing benefit for letter H	20 minutes transfer: random (mean ± SD = 23.8 ± 5.7 s) blocked (mean ± SD = 28.3 ± 6.2 s) ($F_{1,46}$ = 4.4, MSE = 165.6)

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Continued Ste-Marie et al., 2004 [25] Experiment 2 *(Continued)*	Total: 50 48 included in analysis		Tasks themselves	Tasks themselves (either in blocked or random order)	Same tasks (same letters but in cursive script continuously)		*(Continued)* Random: significant writing benefit for letter A 24 hours retention: no significant group differences	*(Continued)* 24 hours transfer: random (mean ± SD = 25.9 ± 4.8 s) > blocked (mean ± SD = 28.2 ± 5.5 s) $F_{1,46}$ = 8.7, MSE = 91.5)
	Blocked: 24 Random: 24		Scores (scoring of the handwritten symbols)	Scores (scoring of the handwritten symbols)	Time taken to write the word			
Ste-Marie et al., 2004 [25] Experiment 3	Children, recruited from 5 schools (Montessori, public and Catholic school) (mean age = 6.42 yrs., SD = 0.40 yrs.)	Handwriting tasks (one-toone-situation, two experiment-ers, each testing a child)	1 session (35 minutes) 72 trials	After 20 minutes after 24 hours	After 20 minutes after 24 hours (after the retention test)	No significant group difference	20 minutes retention: no significant group difference blocked (mean ± SD = 3.60 ± 0.97) = random (mean ± SD = 4.30 ± 0.83) (p = 0.10) 24 hours retention: no significant group difference	20 minutes transfer: random (mean ± SD = 22.91 ± 7.22 s) > blocked (mean ± SD = 28.41 ± 9.46 s) ($F_{1,66}$ = 7.2, MSE=69.9, p = not reported) 24 hours transfer: random (mean ± SD = 20.13 ± 6.43 s) > blocked (mean ± SD = 24.76 ± 9.38 s) ($F_{1,66}$ = 5.6, MSE = 63.6, p = not reported)
	Total: 78 68 included in analysis		Tasks themselves	Tasks themselves (either in blocked or random order	Same tasks (same letters but in cursive script continuously)			
	Blocked: NA Random: NA		Scores (scoring of the handwritten symbols	Scores (scoring of the handwritten symbols)	Time taken to write the word			
Wegman, 1999 [44]	Female 4th grade students (mean age = NA, SD = NA)	Ball rolling, racket striking, ball kicking	1 session (45 minutes) 39 trials (13 trials / task)	After 3 weeks	NA	Ball rolling: blocked (mean ± SD = 64.41 ± 17.75) > random (mean ± SD = 48.94 ± 17.76) ($F_{2,53}$ = 4.63, p < 0.05)	Ball rolling: no significant group difference: blocked (mean ± SD = 58.12 ± 16.21) = random (mean ± SD = 55.00 ± 18.85) ($F_{2,49}$ = 0.34, p > 0.05)	NA

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Wegman, 1999 [44] *(Continued)*	Total: 54		Tasks themselves	Tasks themselves	NA	*(Continued)* Racket striking: blocked (mean ± SD = 66.47 ± 6.31) > random (mean ± SD = 59.16 ± 11.78) ($F_{2,53}$ = 3.22, p < 0.05) Ball kicking: blocked (mean ± SD = 63.75 ± 6.19) > random (mean ± SD = 52.63 ± 20.77) ($F_{2,52}$ = 3.35, p < 0.05)	*(Continued)* Racket striking: random (mean ± SD = 74.16 ± 11.91) > blocked (mean ± SD = 54.41 ± 20.45) ($F_{2,49}$ = 4.73, p < 0.05) posttest: $F_{1,31}$ = 7.41, p < 0.05) Ball kicking: no significant goup difference: blocked (mean ± SD = 50.62 ± 21.74) = random (mean ± SD = 48.68 ± 20.53) ($F_{2,49}$ = 0.24, p > 0.05)	
	Repetitions: NA Random: NA Combined: NA		Scores (accuracy)	Scores (accuracy)	NA			
Wrisberg & Mead, 1983 [45]	Right handed children (mean age = 7 yrs., 2 months, SD = NA)	Anticipation timing task (visual tracking of a moving light, different speeds)	2 days 1 session / day 48 trials / day	NA	After one day 20 trials	No significant group effect ($F_{3,40}$ = 0.37, p > 0.05)	NA	Mean absolute error: Slow velocity transfer: blocked (mean ± SD = 97 ± 56) > random (mean ± SD = 146 ± 63) ($F_{4,50}$ = 3.63, p < 0.05) Fast velocity transfer: no signif-icant group difference Mean variable error: no significant group differences Constant error: significant type of training x blocks interaction ($F_{4,50}$ = 2.92, p < 0.05) Block 2: late responding significantly different: Varied-random-speed: mean ± SD = 41 ± 83 ms (too late), Varied-blocked-speed: mean ± SD =- 19 ± 41 ms (too early)
	Total: 60		Task itself	NA	Task itself, 2 different speeds			
	Slow speed: 12* Fast speed: 12* Varied-random: 12* Varied-blocked: 12* Control: 12* ***Equal number of females and males***							

(Continued)

Table 2.1. Continued

Study	Description		Acquisition	Retention	Transfer	Results		
	Participants	*Task/skill*	*Timeframe*	*Timeframe*	*Timeframe*	*Acquisition*	*Retention*	*Transfer*
	n total		*Ourtcome measures*	*Ourtcome measures*	*Ourtcome measures*			
	n group		*Outcome parameters*	*Outcome parameters*	*Outcome parameters*			
Zetou et al., 2007 [46]	Female, unskilled volleyball players (mean age = 12.4 yrs., SD = 1.2)	Volleyball skills (set, pass, service)	10 weeks 2 sessions / week (75 minutes and 72 trials each)	After 2 weeks	NA	Set: no significant group difference (main effect for group: $F_{1,24}$ = 0.11, p = 0.74) Pass: no significant group difference (main effect for group: $F_{1,24}$ = 0.04, p = 0.84) Service: no significant group difference (main effect for group: $F_{1,24}$ = 0.04, p = 0.85)	Set: no significant group difference (main effect for group: $F_{1,24}$ = 0.11, p = 0.74) Pass: no significant group difference (main effect for group: $F_{1,24}$ = 0.04, p = 0.84) Service: no significant group difference (main effect for group: $F_{1,24}$ = 0.04, p = 0.85)	NA
	Total: 26		Tasks themselves	Tasks themselves	NA			
	Low interference: 13 High interference: 13		Scores (ability to play the ball, hitting target)	Scores (ability to play the ball, hitting target)	NA			

Results: between-groups-effects (blocked vs. random practice) were evaluated. If additional groups (e.g. serial, repetitive) were evaluated, these results were not considered. Abbreviations: ACE = absolute constant error; CA = chronological age; CI = contextual interference; CP = cerebral palsy; DS = Down's Syndrome; IDD = intellectual developmental disability; LD = learning disabilities; m = meters; MA = mental age; ms = milliseconds; MSE = mean square error; n = number; NA = not applicable; SD = standard deviation; TD = typically developing; yr./yrs. = year/years; > meaning "better than".
* n per group is not mentioned in the paper, the information was given by the authors answering to our e-mail.

Table 2.2. Levels of evidence and conduct quality

Study	Evidence level	Quality							
		Conduct questions							
		Summary	1	2	3	4	5	6	7
Bertollo et al., 2010 [39]	II	0/7	no	no	no	no	no	no	no
Bortoli et al., 1992 [37]	II	0/7	no	no	no	no	no	no	no
Broadbent et al., 2015 [30]	III	2/7	no	no	yes	no	no	yes	no
Del Rey et al., 1983 [31]	III	0/7	no	no	no	no	no	no	no
Duff et al., 2003 [28]	II	0/7	no	no	no	no	no	no	no
Edwards et al., 1986 [6]	II	0/7	no	no	no	no	no	no	no
Fialho et al., 2006 [32]	III	0/7	no	no	no	no	no	no	no
French et al., 1990 [29]	I	0/7	no	no	no	no	no	no	no
Gophna et al., 2007 [26]	II	0/7	no	no	no	no	no	no	no
Granda Vera & Montilla, 2003 [33]	III	0/7	no	no	no	no	no	no	no
Green et al., 1995 [41]	II	0/7	no	no	no	no	no	no	no
Jarus & Goreover, 1999 [34]	III	0/7	no	no	no	no	no	no	no
Jarus & Gutman, 2001 [14]	III	0/7	no	no	no	no	no	no	no
Jones & French, 2007 [38]	II	1/7	no	no	no	no	no	yes	no
Meira & Tani, 2003 [42]	II	0/7	no	no	no	no	no	no	no
Painter et al., 1994 [27]	II	0/7	no	no	no	no	no	no	no
Perez et al., 2005 [40]	II	0/7	no	no	no	no	no	no	no
Pigott & Shapiro, 1984 [7]	III	0/7	no	no	no	no	no	no	no
Pollock & Lee, 1997 [43]	II	0/7	no	no	no	no	no	no	no
Saemi et al., 2012 [44]	II	0/7	no	no	no	no	no	no	no
Stambaugh, 2011 [45]	II	0/7	no	no	no	no	no	no	no
Ste-Marie et al., 2004 [25] Experiment 1	II	1/7	no	no	no	yes	no	no	no
Ste-Marie et al., 2004 [25] Experiment 2	II	2/7	no	no	no	yes	no	yes	no
Ste-Marie et al., 2004 [25] Experiment 3	II	2/7	no	no	no	yes	no	yes	no
Wegman, 1999 [36]	II	0/7	no	no	no	no	no	no	no
Wrisberg & Mead, 1983 [35]	III	0/7	no	no	no	no	no	no	no
Zetou et al., 2007 [46]	II	0/7	no	no	no	no	no	no	no

Evidence levels and scoring of the conduct questions of all the included studies. Evidence levels: level I = randomised controlled trials (sample size > 100); level II randomised controlled trials (sample size < 100); level III: controlled cohort studies; level IV: case series; level V: expert opinions [15]. Conduct questions: 1) inclusion and exclusion criteria, 2) description of and adherence to the intervention, 3) validity and reliability of outcome measures, 4) masking of the participants and assessors, 5) statistical analysis, 6) dropouts, 7) controlling for confounding variables [15].

Table 2.3. Best evidence synthesis according to tasks, learning level and practice order

Area		Task	Study	Evidence synthesis per study			Evidence synthesis summary		
				Acquisition	Retention	Transfer	Acquisition	Retention	Transfer
NLT	FMT	Mirror tracing task	Gophna et al., 2007 [26]	NS/NR	NS	NS	-	-	-
		Playing clarinet	Stambaugh, 2011 [45]	IC	IC	NS	X	X	-
		Handwriting skills	Ste-Marie et al., 2004 [25] Experiment 1	NS	R	NA	X	X	**R
			Ste-Marie et al., 2004 [25] Experiment 2	IC	IC	R			
			Ste-Marie et al., 2004 [25] Experiment 3	NS	NS	R			
	GMT	Throwing beanbags	Jarus & Goreover, 1999 [34]	IC	IC	IC	X	X	X
			Jarus & Gutman, 2001 [14]	IC	NS	NS			
			Painter et al., 1994 [27]	NS/NR	NS^/NR	NA			
			Pigott & Shapiro, 1984 [7]	NS	NA	NS			
		Throwing different balls	Granda Vera & Montilla, 2003 [33]	IC	R	R	X	*R	*R
		Throwing tennis balls	Saemi et al., 2012 [44]	NS	NS	NA	-	-	NA
		Volleyball	Bortoli et al., 1992 [37]	NS	NS	IC	-	-	X
			Fialho et al., 2006 [32]	NS/NR	NA	IC			
			French et al., 1990 [29]	NS	NS	NA			
			Jones & French, 2007 [38]	NS	NS	NA			
			Meira & Tani, 2003 [42]	NS	NA	NS			
			Zetou et al., 2007 [46]	NS	NS	NA			
		Hitting different balls with different rackets	Green et al., 1995 [41]	NS	NA	IC	-	NA	X
		Lifting unfamiliar objects	Duff et al., 2003 [28]	NR	NR	NA	-	-	NA
		Dance step sequence	Bertollo et al., 2010 [39]	B	NS	NA	*B	-	NA
		Propelling task	Pollock & Lee, 1997 [43]	NS	NR	NR	-	-	-
		Tennis	Broadbent et al., 2015 [30]	NS	R	IC	-	*R	X
		Ball rolling, striking, kicking	Wegman, 1999 [36]	B	IC	NA	*B	X	NA
LT	LMT	Anticipation timing task	Del Rey et al., 1983 [31]	B	NA	NS	X	NA	X
			Edwards et al., 1986 [6]	NR	NA	IC			
			Wrisberg & Mead, 1983 [35]	NS	NA	IC			
		Positioning motor task	Perez et al., 2005 [40]	B	IC	NA	*B	X	NA

Abbreviations: NLT = Non-laboratory tasks; LT = Laboratory tasks; FMT = Fine-motor tasks GMT = Gross motor tasks; B = significant, favouring blocked order; IC = inconsistent; NA = not applicable, no study evaluated the according aspect; NR= not reported; NS = not significant; R = significant, favouring random order. Evaluation of the studies: Results of the single studies were evaluated taking in account the typically developing children and all parameters and tasks into account. Results with ≥ 75% of the comparisons favouring one practice order were evaluated as consistent evidence within one study. Evaluation of the tasks: Results of the according studies were merged if ≥ 75% of the studies of one task showed the same result, evidence was rated as consistent.

Table 2.3. Continued

Strength of the evidence (adapted from Tulder et al. [24]): *** = Strong- consistent findings among multiple high quality randomised controlled trials (RCTs); ** = Moderate- consistent findings among multiple low quality RCTs and/or controlled clinical trials (CCTs) and/or high one high quality RCT ; * = Limited - one low quality RCT and/or CCT; X = Conflicting- inconsistent findings among multiple trials (RCTs and/or CCTs), inconsistent findings among different parameters within one trial (if only one trial is available) ;- = No evidence from trials- no RCTs or CCTs.

DISCUSSION

We investigated the evidence of contextual interference in children with congenital or acquired brain injuries and typically developing children. Only one study included children with brain lesions. The methodological quality of the studies was low and the risk of bias high, which makes it difficult to formulate recommendations whether children with brain lesions or typically developing children would profit more from a blocked or random approach.

Contextual interference in children with hemiplegic cerebral palsy

The one study examining children with hemiplegic CP (n=18, mean age 10 years, SD 1.8) also included a group of age-matched typically developing peers (n=18, mean age 10.4 years, SD 1.7 years) [28].The study consisted of two experiments. In the first experiment, children lifted various known objects while the vertical lifting load force was measured. The second experiment investigated the contextual interference effect. The participants lifted three novel objects with varying weights 27 times. One group did this in blocked order, the other group in random order. Retention was tested immediately after and 24 hours after the practice phase. While during acquisition blocked practice resulted in better differentiation of force rates between the different objects, there was no difference during the retention trials between the practice groups. Based on these two experiments the authors concluded that children with hemiplegic CP have an internal picture of the weight of familiar objects, that they can learn and retain to provide the right amount of force when lifting objects with unknown weights, but that the amount of practice rather than the practice order is essential for this learning process [28]. A conclusion about the contextual interference effect in children with CP is difficult, though, because this was the only study we found and it had some qualitative shortcomings. In the methodological quality assessment, this study received 0 of 7 points (Table 2.2) and the risk of bias was high. The main reasons for our low rating of this study were the lack of information about the control group at baseline (only the means of the whole groups are reported without a measure of variation), the missing information about the adherence, the psychometric properties of the assessments were not reported, it was unclear whether assessors were masked, the lack of a power calculation, and the number of drop-outs were not reported.

Contextual interference in typically developing children

When considering contextual interference studies with typically developing children, a conclusion also remains unclear. Although the best evidence synthesis showed limited to moderate support of the contextual interference effect for some of the tasks (favouring random practice for better retention and transfer), in the majority of the tasks no evidence (acquisition: n=7/15, retention: 6/13, transfer: n=3/10) or conflicting evidence (acquisition: n=5/15, retention: n=5/13, transfer: n=5/10) was found (Table 2.3). Besides the low methodological quality, several factors could have affected the contextual interference effect contributing to the inconclusiveness of some results [2].

The influence of types of skills and variations

One of these factors might be the kind of skill and its variations that were studied. For example, Magill and Hall already discussed that the generalisability of the contextual interference effect could be influenced by task characteristics like laboratory tasks, such as coincident anticipation timing tasks, versus motor skill performance outside the laboratory or non-laboratory tasks, such as throwing beanbags [5].

In our review, we found four studies that investigated laboratory tasks, namely anticipation timing tasks [6, 31, 35], and a positioning motor task [40]. The other studies investigated non-laboratory tasks: six examined volleyball skills [29, 32, 37, 38, 42, 46], four beanbag throwing [7, 14, 27, 34], and three experiments investigated handwriting skills [25]. Further tasks were dance step sequences [39], tennis skills [30], lifting unfamiliar objects [28], mirror tracing tasks [26], throwing different balls [33], hitting different balls with different rackets [41], rolling, striking and hitting balls [36], a propelling task (Crokinole game) [43], throwing tennis balls [44], and playing the clarinet [45].

From a therapeutic point of view, this change of interest from laboratory to non-laboratory tasks is desired. While learning a laboratory task can reflect the capacity of the child, i.e. what a person with a health condition actually is able to do, learning non-laboratory tasks may better resemble daily life activities, i.e. performance (what a person does in his/her usual environment, e.g. skills or tasks needed for self-care, leisure activities, school or work), as described by the World Health Organisation's International Classification of Functioning, Disability and Health (ICF) [47]. Practicing non-laboratory tasks might improve the translation to other daily life relevant tasks, as these tasks might appear more natural and are probably more frequently occurring in the child's daily routines than laboratory tasks.

In healthy adults, the evidence is mixed when practicing laboratory tasks, but practicing non-laboratory tasks supports the contextual interference effect [5]. In our review, the best evidence synthesis shows limited to moderate support for the contextual interference effect in five non-laboratory and one laboratory task

(Table 2.3). However, only in a "throwing different balls" [33] task, we found the contextual interference effect for both the retention and transfer. In summary, a clear differentiation about how laboratory or non-laboratory tasks support the contextual interference effect in children cannot be determined.

The influence of experience, age and task difficulty

In adults, it is suggested that the learner needs to have an idea of the movement or some initial experience to profit from variations of the practice schedule, but the relation between age and the contextual interference remains unclear [5]. In children, Jarus and Goreover observed a difference between three age groups (5 years, 7 years, and 11 years) who practiced beanbag throwing [34]. In general, older children performed the bean bag throwing task better. While the group of 7-year-old children acquired and retained better during blocked practice, there was no difference between the practice groups during the transfer [34]. Hence, for this task, blocked practice might be more beneficial for this age group. While the authors argued that this finding could be explained by the low experience level and the young age of the 7-year-old children, these observations were not made for the 5-year-old children, which is not in line with the author's explanation. Furthermore, Pollock and Lee compared the learning of propelling a small wooden disk with the middle finger (an adaptation of the Crokinole game) between children and adults [43]. They could reconstruct the contextual interference effect in adults and also in children during transfer and retention, but the children showed no difference in task acquisition between the blocked and random groups [43]. Apparently, this pattern has been observed in adults practicing a difficult task and can be explained with a benefit of blocked practice during acquisition of easy tasks only [2, 48].

If the difficulty level of a task seems to influence only the acquisition but not the retention and transfer in adults, the random practice order can be recommended in adults, regardless whether the task is simple or difficult. In typically developing children, though, we cannot make such a recommendation, because the evidence is unclear as experience, age, and task difficulty intermingle with each other. It becomes even more complicated when trying to generalise the effects of experience, age, and task difficulty on the contextual interference effect to children with congenital or acquired brain lesions. For example, the question of whether children in neurorehabilitation are novices or whether they are experienced has to be considered carefully. Given that children with congenital brain lesions or acquired brain lesions in a chronic state have had therapy for most of their lives or for a long time, respectively, they could be considered an expert group. Children with a (sub-) acute acquired brain injury most likely could, on the one hand, be regarded as novices when it comes to relearning motor activities of daily life with their impaired sensorimotor and cognitive systems. On the other hand, they might have performed all these activities independently before they experienced the brain injury which

puts them on another starting point compared to children with congenital brain lesions. All these aspects need to be incorporated when considering the contextual interference effect in paediatric neurorehabilitation.

Can knowledge about related populations be transferred to children with brain lesions?

The evidence about the contextual interference in typically developing children is limited, yet, more extensive compared to the evidence available for children with brain lesions. In adults with brain lesions, the results are not clearly supporting the contextual interference effect. In adults with chronic stroke, for example, the typical contextual interference effect could not be reconstructed when performing three specific movements (wrist/finger extension, elbow joint extension, and shoulder joint abduction) combined with active neuromuscular stimulation [49]. When practicing a task that was designed to approximate the steps needed take a coffee cup out of the cupboard and put it on the table, the random practice outperformed the blocked group in stroke patients [50]. Schweighofer et al. [51] concluded that these conflicting results might be due to the lack of separation of the patients between high and low working memory capabilities. In their study, individuals with stroke with normal visuospatial working memory retained visuomotor skills better when practicing in random order compared to blocked order, while in participants with low visuospatial working memory retention performance did not differ between practice groups [51]. It seems that also in adults with brain injury results might be influenced by other factors additional to the practice order.

We assume that results obtained in typically developing children (compared to healthy adults) could be better generalised to children with brain lesions. Nevertheless, we should be cautious, because, firstly, the physical requirements are different: damaged sensory pathways and structures involved in processing sensory information, such as found in children with brain lesions, reduce the ability to detect errors and consequently impair motor learning [52, 53]. Secondly, learning a motor task with a damaged brain is likely different from learning with an intact, typically developing brain. This stands in contrast with previous observations in adults with a unilateral stroke which suggested that the stroke affected the control and execution, but not the learning of motor skills per se [54]. However, as this issue is under debate, there is still no definite agreement which brain regions and processes are involved in learning and how the learning processes are executed (e.g. [55–57]. Furthermore, the individual lesion areas in combination with many other factors make every patient and their learning abilities and strategies unique. Therefore, further research is needed to understand the relation of pathological changes and motor learning disorders [58]. Thirdly, it has been shown that physiotherapists perceive primary impairments (e.g. muscle tone, movement patterns) and secondary outcomes (e.g. range of motion, joint alignment, muscle strength), as well as personal factors

(e.g. motivation) and environmental factors (e.g. support and expectations from the family) as important factors influencing the acquisition of motor abilities in children with CP [59]. These factors could slow down or even hinder learning in children with brain lesions compared to typically developing children.

The methodological quality of the studies

A reconstruction of the methodological approach was challenging in many studies. It was often unclear whether certain methodological aspects were poorly performed or just poorly described. This influenced our assessment of bias and quality. For example, the psychometric properties of the applied outcome measures were unknown or not reported. Several studies mentioned reliability evaluations of their measures, while information on validity or absolute measurements errors (such as the standard error of measurement) was missing. Also, the description of the appropriate statistics and power calculation (both are needed to score a "yes") was missing. Only one study mentioned a power analysis but did not present it [34].

The oldest papers we included in this review were published in the 1980ies (Figure 2.2). The Standardised Reporting of Trials (SORT) Statement [60] was published in 1994. Before that checklists for interventional trials were not available. This might partly explain why the older studies did not report all aspects systematically and were therewith rated low in the quality evaluation.

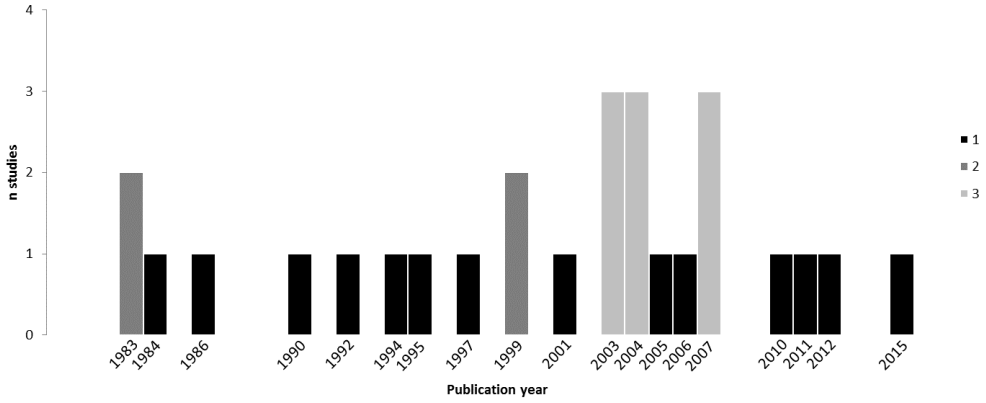


Fig. 2.2. The distribution of the publication years of the articles included in this systematic review.

Limitations

There are some limitations of this systematic review which need to be mentioned. Our literature search was limited to seven databases and restricted to published articles only. Grey literature was not considered. We excluded studies which performed a proceeding familiarisation phase prior to the actual practice phase, whether or not this affects the learning remains to be discussed. For the best evidence synthesis, we

did not subdivide the study results according to long- or short-term learning phases because there were not enough comparable studies to build subgroups. This asks for a cautious interpretation of the results.

Recommendations for future research

We expect that the contextual interference effect in children with brain lesions can influence rehabilitation outcomes. Therefore, we would recommend to design such studies and include these particular patient groups. A careful selection of the motor task to be studied is crucial: it should be clinically relevant and motivating for the child to perform and it should provide objective parameters to quantify the retention or transfer of the task or skill particularly, as these are most relevant for the child after discharge from rehabilitation. The study should be designed and its results reported in accordance with the various internationally accepted checklists to ensure high study quality and low bias.

CONCLUSION

To recapitulate, there is a persistent demand for increasing our knowledge about the contextual interference effect in children, especially, in children with brain lesions, as the number of existing studies is small, and the methodological quality of the studies is low. For some tasks, we found limited evidence supporting the contextual interference effect in typically developing children. However, we would be cautious in generalising these results to children with brain lesions. To improve movement or sports programmes in typically developing children and advance rehabilitation programmes for children with brain lesions, there is an emerging need to increase our knowledge of the contextual interference effect in these populations.

ACKNOWLEDGMENTS

We would like to thank all the authors, who responded to our e-mails and providing us with more information about their studies. A very special thank you goes to Yvonne Perathoner and Martina Hediger, the librarians of the library of the Children's University Hospital Zurich, who supported our search and ordered copies of all the papers we asked for.

AUTHOR CONTRIBUTIONS

Conceptualization: Judith V Graser
Formal analysis: Judith V Graser
Funding acquisition: Hubertus JA van Hedel
Investigation: Judith V Graser, Caroline HG Bastiaenen, Hubertus JA van Hedel

Methodology: Judith V Graser, Caroline HG Bastiaenen, Hubertus JA van Hedel
Project administration: Judith V Graser
Resources: Hubertus JA van Hedel
Validation: Judith V Graser, Caroline HG Bastiaenen, Hubertus JA van Hedel
Visualization: Judith V Graser
Writing – original draft: Judith V Graser
Writing – review & editing: Judith V Graser, Caroline HG Bastiaenen, Hubertus JA van Hedel

SUPPORTING INFORMATION

S2.1 Table. Example of the search strategy used for the primary search on PsycINFO.

On PsycINFO we conducted 4 searches with different filters and combined them.

Basic search term (for all the 4 searches): ('paediatrics' OR 'pediatrics' OR 'children' OR 'child' OR 'childhood' OR 'adolescent' OR 'adolescence' OR 'student' OR 'elementary' OR 'high school' OR 'youths') AND ('motor learning' OR 'skill learning') AND ('contextual interference' OR 'practice order' OR 'random' OR 'blocked') AND ('performance' OR 'acquisition' OR 'retention' OR 'transfer' OR 'generalisation' OR 'generalization')
Filters: Filter for search #1: childhood (birth-12 yrs.) Filter for search #2: school age (6-12 yrs.) Filter for search #3: preschool age (2-5 yrs.) Filter for search #4: adolescence (13-17 yrs.)
Combined search: #1 OR #2 OR #3 OR #4

Table 2.2. Levels of evidence

Level	Intervention/Group studies	Single subject design studies
I	Systematic review of randomised controlled trials (RCTs) Large RCT (with narrow confidence interval) (n>100)	Randomised controlled N-of 1 (RCT), alternating treatment design (ATD), and concurrent or non-concurrent multiple baseline designs (MBDs); generalisability if the ATD is replicated across three or more subjects and the MB consists of a minimum of three subjects, behaviours, or settings. These designs can provide causal inferences.
II	Smaller RCTs (with wider confidence intervals) (n<100) Systematic reviews of cohort studies Outcome research (very large ecologic studies)	Non-randomised, controlled, concurrent MBD; generalisability if design consists of a minimum of tree subjects, behaviours, or settings. Limited causal inferences.
III	Cohort studies (must have concurrent control group) Systematic reviews of case control studies	Non-randomised, non-concurrent, controlled MBD; generalisability if design consists of a minimum of three subjects, behaviours or settings. Limited causal inferences.
IV	Case series Cohort studies without concurrent control group (e.g. with historical control group) Case-control study	Non-randomised, controlled SSRDs with at least three phrases (ABA, ABAB, BAB, etc.); generalisability if replicated across three or more different subjects. Only hints at causal inferences.
V	Expert opinion Case study or report Bench research Expert opinion based on theory or physiological research Common sense/anecdotes	Non-randomised controlled AB SSRD; generalisability if replicated across three or more different subjects. Suggests causal inferences allowing for testing of ideas.

Levels of evidence in intervention or group studies and single-case design studies according to the American Academy for Cerebral Palsy and Developmental Medicine (AACPDM) [15].

Abbreviations: ATD = alternating treatment design; MB = multiple baseline; MBD = multiple baseline design; n = number; RCT = randomised controlled trial; SSRD = single subject research design.

S2.3 Table. Conduct questions

No	Question
1	Were inclusion and exclusion criteria of the study population well described and followed?
2	Was the intervention well described and was there adherence to the intervention assignment? (For 2 group-designs, was the control exposure also well described?) Both parts of the question need to be met to score “yes”.
3	Were all the measures used clearly described, valid and reliable for measuring the outcomes of interest?
4	Was the outcome assessor unaware of the intervention status of the participants (i.e. were the assessors masked?)
5	Did the authors conduct and report appropriate statistical evaluation including power calculations? Both parts of the question need to be met to score «yes».
6	Were dropouts/loss to follow-up reported as less than 20%? For 2-group designs, was dropout balanced?
7	Considering the potential within the study design, were appropriate methods for controlling confounding variables and limiting potential biases used?

Questions to evaluate the methodology of an intervention study according to the American Academy for Cerebral Palsy and Developmental Medicine [15].

S2.4 Table. The Cochrane Collaboration's tool for assessing the risk of bias

Domain	Review authors' judgement
Selection bias	
Random sequence generation	Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence
Allocation concealment	Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.
Performance bias	
Blinding of participants and personnel	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study. *Assessments should be made for each main outcome (or class of outcomes).*
Detection bias	
Blinding of outcome assessment	Detection bias due to knowledge of the allocated interventions by outcome assessors. *Assessments should be made for each main outcome (or class of outcomes).*
Attrition bias.	
Incomplete outcome data	Attrition bias due to amount, nature or handling of incomplete outcome data. *Assessments should be made for each main outcome (or class of outcomes).*
Reporting bias	
Selective reporting	Reporting bias due to selective outcome reporting.
Other bias	
Other sources of bias	Bias due to problems not covered elsewhere in the table.

The risk of bias tool, presented by the Cochrane handbook for systematic reviews of interventions [16].

REFERENCES

1. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: The definition and classification of cerebral palsy April 2006. Dev Med Child Neurol. 2007;49 SUPPL.109:8–14.
2. Lee TD, Wulf G, Schmidt RA. Contextual interference in motor learning: Dissociated effects due to the nature of task variations. Q J Exp Psychol Sect A Hum Exp Psychol. 1992;44:627–44.
3. Battig W. Intratask interference as a source of facilitation on transfer and retention. In: Voss J, editor. Topics in learning and performance. New York: Academic Press; 1972. p. 131–59.
4. Hodges NJ, Lohse KR, Wilson A, Lim SB, Mulligan D. Exploring the dynamic nature of contextual interference: Previous experience affects current practice but not learning. J Mot Behav. 2014;46:455–67.
5. Magill RA, Hall KG. A review of the contextual interference effect in motor skill acquisition. Hum Mov Sci. 1990;9:241–89.
6. Edwards JM, Elliot D, Lee TD. Contextual interference effects during skill acquisition and transfer in Down's Syndrome adolescents. Adapt Phys Act Q. 1986;3:250–8.
7. Pigott RE, Shapiro DC. Motor schema: The structure of the variability session. Res Q Exerc Sport. 1984;55:41–5.
8. Rey P Del, Whitehurst M, Wughalter E, Barnwell J. Contextual interference and experience in acquisition and transfer. Percept Mot Skills. 1983;57:241–2.
9. Brady F. The contextual interference effect and sport skills. Percept Mot Skills. 2008;106:461–72.
10. Brady F. A theoretical and empirical review of the contextual interference effect and the learning of motor skills. Quest. 1998;50:266–93.
11. Merbah S, Meulemans T. Learning a motor skill: Effects of blocked versus random practice- A review. Psychol Belg. 2011;51:15–48.
12. Barreiros J, Figueiredo T, Godinho M. The contextual interference effect in applied settings. Eur Phys Educ Rev. 2007;13:195–208.
13. Kitago T, Krakauer JW. Motor learning principles for neurorehabilitation. In: Barnes MP, Good DC, editors. Handbook of Clinical Neurology. Amsterdam: Elsevier; 2013. p. 93–103.
14. Jarus T, Gutman T. Effects of cognitive processes and task complexity on acquisition, retention, and transfer of motor skills. Can J Occup Ther. 2001;68:280–9.

15. Darrah J, Hickman R, O 'Donnell M, Vogtle L, Wiart L. AACPDM Methodology to develop systematic reviews of treatment interventions (revision 1.2). 2008. https://www.aacpdm.org/UserFiles/file/systematic-review-methodology.pdf.

16. Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011.

17. Harlow HF. The formation of learning sets. Psychol Rev. 1948;56:51–65.

18. Seidler RD. Multiple motor learning experiences enhance motor adaptability. J Cogn Neurosci. 2004;16:65–73.

19. Braun DA, Aertsen A, Wolpert DM. Motor task variation induces structural learning. Curr Biol. 2009;19:352–7.

20. Kattner F, Cochrane A, Cox CR, Gorman TE, Green CS. Perceptual learning generalization from sequential perceptual training as a change in learning rate. Curr Biol. 2017;27:840–6.

21. Bejjanki VR, Zhang R, Li R, Pouget A, Green CS, Lu Z-L, et al. Action video game play facilitates the development of better perceptual templates. Proc Natl Acad Sci U S A. 2014;111:16961–6.

22. Brown A, Kane M. Preschool children can learn to transfer: Learning to learn and learning from example. Cogn Psychol. 1988;20:493–523.

23. Giboin L-S, Gruber M, Kramer A. Motor learning of a dynamic balance task: Influence of lower limb power and prior balance practice. J Sci Med Sport. 2019;22:101–5.

24. Van Tulder M, Furlan A, Bombardier C. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. Spine (Phila Pa 1976). 2003;28:1290–9.

25. Ste-Marie DM, Clark SE, Findlay LC, Latimer AE. High levels of contextual interference enhance handwriting skill acquisition. J Mot Behav. 2004;36:115–26.

26. Gophna S, Jarus T, Weintraub N. Effects of contextual interference on acquisition, retention and transfer of motor skill of children with and without learning disabilities. Isr J Occup Ther. 2007;16:89–108.

27. Painter MA, Inman KB, Vincent WJ. Contextual interference effects in the acquisition and retention of motor tasks by individuals with mild mental handicaps. Adapt Phys Act Q. 1994;11:383–95.

28. Duff S V, Gordon AM. Learning of grasp control in children with hemiplegic cerebral palsy. Dev Med Child Neurol. 2003;45:746–57.

29. French KE, Rink JE, Werner PH. Effects of contextual interference on retention of three volleyball skills. Percept Mot Skills. 1990;71:179–86.

30. Broadbent DP, Causer J, Ford PR, Williams AM. Contextual interference effect on perceptual–cognitive skills training. Med Sci Sport Exerc. 2015;47:1243–50.

31. Del Rey P, Whitehurst M, Wood J. Effects of experience and contextual interference on learning and transfer by boys and girls. Percept Mot Skills. 1983;56:581–2.

32. Fialho JVAP, Benda RN, Ugrinowitsch H. The contextual interference effect in a serve skill acquisition with experienced volleyball players. J Hum Mov Stud. 2006;50:65–77.

33. Granda Vera J, Montilla MM. Practice schedule and acquisition, retention, and transfer of a throwing task in 6-yr.-old children. Percept Mot Skills. 2003;96 3 Pt 1:1015–24.

34. Jarus T, Goreover Y. Effects of contextual interference and age on acquisition, retention, and transfer of motor skill. Percept Mot Skills. 1999;88:437–47.

35. Wrisberg C, Mead BJ. Developing coincident timing skill in children: A comparison of training methods. Res Q Exerc Sport. 1983;54:67–74.

36. Wegman E. Contextual interference effects on the acquisition and retention of fundamental motor skills. Percept Mot Skills. 1999;88:182–7.

37. Bortoli L, Robazza C, Durigon V, Carra C. Effects of contextual interference on learning technical sports skills. Percept Mot Skills. 1992;75:555–62.

38. Jones LL, French KE. Effects of contextual interference on acquisition and retention of three volleyball skills. Percept Mot Skills. 2007;105:883–90.

39. Bertollo M, Berchicci M, Carraro A, Comani S, Robazza C. Blocked and random practice organization in the learning of rhythmic dance step sequences. Percept Mot Skills. 2010;110:77–84.

40. Perez CR, Meira CM, Tani G. Does the contextual interference effect last over extended transfer trials? Percept Mot Skills. 2005;100:58–60.

41. Green DP, Whitehead J, Sugden DA. Practice variability and transfer of a racket skill. Percept Mot Skills. 1995;81:1275–81.

42. Meira CMJ, Tani G. Contextual interference effects assessed by extended transfer trials in the acquisition of the volleyball serve. J Hum Mov Stud. 2003;45:449–68.

43. Pollock BJ, Lee TD. Dissociated contextual interference effects in children and adults. Percept Mot Skills. 1997;84 3 Pt 1:851–8.

44. Saemi E, Porter JM, Varzaneh AG, Zarghami M, Shafinia P. Practicing along the contextual interference continuum: A comparison of three practice schedules in an elementary physical education setting. Kinesiology. 2012;44:191–8.

45. Stambaugh LA. When repetition isn't the best practice strategy: Effects of blocked and random practice schedules. J Res Music Educ. 2011;58:368–83.

46. Zetou E, Michalopoulou M, Giazitzi K, Kioumourtzoglou E. Contextual interference effect in learning volleyball skills. Percept Mot Skills. 2007;104:995–1004

47. World Health Organization. International classification of functioning, disability and health - children & youth version. Geneva, CH: WHO Press, World Health Organization; 2007.

48. Tsutsui S, Lee TD, Hodges NJ. Contextual interference in learning new patterns of bimanual coordination. J Mot Behav. 1998;30:151–7.

49. Cauraugh JH, Kim SB. Stroke motor recovery: Active neuromuscular stimulation and repetitive practice schedules. J Neurol Neurosurg Psychiatry. 2003;74:1562–6.

50. Hanlon RE. Motor learning following unilateral stroke. Arch Phys Med Rehabil. 1996;77:811–5.

51. Schweighofer N, Lee J, Goh H, Choi Y, Kim SS, Stewart JC, et al. Mechanisms of the contextual interference effect in individuals poststroke. J Neurophysiol. 2011;106:2632–41.

52. Seidler RD, Kwak Y, Fling BW, Bernard JA. Neurocognitive mechanisms of error-based motor learning. Adv Exp Med Biol. 2013;782:39–60.

53. Krakauer JW, Ghilardi M-F, Ghez C. Independent learning of internal models for kinematic and dynamic control of reaching. Nat Neurosci. 1999;2:1026–31.

54. Winstein C. Motor learning after unilateral brain damage. Neuropsychologia. 1999;37:975–87.

55. Gregory E, McCloskey M, Ovans Z, Landau B. Declarative memory and skill-related knowledge: Evidence from a case study of amnesia and implications for theories of memory. Cogn Neuropsychol. 2016;33:220–40.

56. Steinemann NA, Moisello C, Ghilardi MF, Kelly SP. Tracking neural correlates of successful learning over repeated sequence observations. Neuroimage. 2016;137:152–64.

57. McNamee D, Liljeholm M, Zika O, O'Doherty JP. Characterizing the associative content of brain structures involved in habitual and goal-directed actions in humans: A multivariate FMRI study. J Neurosci. 2015;35:3764–71.

58. Ding Y, Yao B, Lai Q, McAllister JP. Impaired motor learning and diffuse axonal damage in motor and visual systems of the rat following traumatic brain injury. Neurol Res. 2001;23:193–202.

48. Tsutsui S, Lee TD, Hodges NJ. Contextual interference in learning new patterns of bimanual coordination. J Mot Behav. 1998;30:151–7.

49. Cauraugh JH, Kim SB. Stroke motor recovery: Active neuromuscular stimulation and repetitive practice schedules. J Neurol Neurosurg Psychiatry. 2003;74:1562–6.

50. Hanlon RE. Motor learning following unilateral stroke. Arch Phys Med Rehabil. 1996;77:811–5.

51. Schweighofer N, Lee J, Goh H, Choi Y, Kim SS, Stewart JC, et al. Mechanisms of the contextual interference effect in individuals poststroke. J Neurophysiol. 2011;106:2632–41.

52. Seidler RD, Kwak Y, Fling BW, Bernard JA. Neurocognitive mechanisms of error-based motor learning. Adv Exp Med Biol. 2013;782:39–60.

53. Krakauer JW, Ghilardi M-F, Ghez C. Independent learning of internal models for kinematic and dynamic control of reaching. Nat Neurosci. 1999;2:1026–31.

54. Winstein C. Motor learning after unilateral brain damage. Neuropsychologia. 1999;37:975–87.

55. Gregory E, McCloskey M, Ovans Z, Landau B. Declarative memory and skill-related knowledge: Evidence from a case study of amnesia and implications for theories of memory. Cogn Neuropsychol. 2016;33:220–40.

56. Steinemann NA, Moisello C, Ghilardi MF, Kelly SP. Tracking neural correlates of successful learning over repeated sequence observations. Neuroimage. 2016;137:152–64.

57. McNamee D, Liljeholm M, Zika O, O'Doherty JP. Characterizing the associative content of brain structures involved in habitual and goal-directed actions in humans: A multivariate FMRI study. J Neurosci. 2015;35:3764–71.

58. Ding Y, Yao B, Lai Q, McAllister JP. Impaired motor learning and diffuse axonal damage in motor and visual systems of the rat following traumatic brain injury. Neurol Res. 2001;23:193–202.

59. Bartlett DJ, Palisano RJ. Physical therapists' perceptions of factors influencing the acquisition of motor abilities of children with cerebral palsy: Implications for clinical reasoning. Phys Ther. 2002;82:237–48.

60. The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. J Am Med Assoc. 1994;272:1926–31.

CHAPTER 3

RELIABILITY OF AND PRACTICAL LESSONS LEARNT FROM ROBOTIC UPPER LIMB ASSESSMENTS IN CHILDREN UNDERGOING NEUROREHABILITATION

Judith V Graser
Laura Prospero
Monica Liesch
Urs Keller
Hubertus JA van Hedel

Submitted

ABSTRACT

Background: In children with disorders of the central nervous system, upper limp function is often impaired. As this affects activities of daily life, assessing upper limb function comprehensively is important for planning and evaluating neurorehabilitative interventions. Measuring with robotic devices increases objectivity and enables to measure functions such as quality of a movement or maximal reachable workspace, which are usually not assessed with conventional assessments. The aim was to evaluate the relative and absolute test-retest reliability of seven assessments provided by the upper limb exoskeleton robot ChARMin.

Methods: Thirty participants (mean age ± standard deviation 12.5 ± 3.3 years, 9 females) with impaired upper limb function due to congenital (n=15) or acquired (n=14) brain injuries (or both, n=1) performed the ChARMin assessments active and passive Range of Motion (aROM, pROM), Strength, Resistance to Passive Movement (RPM), Quality of Movement (QoM), Circle, and Workspace on two measurement sessions, three to seven days apart. Intraclass Correlation Coefficients (ICCs, relative reliability), the Smallest Real Differences (SRD) and SRD as percentage of the grand mean (SRD%, absolute reliability) were calculated for the parameters of each assessment between the two sessions.

Results: Relative reliability ranged from little (e.g. RPM 'resistance against shoulder external rotation': ICC=-0.03, 95% confidence interval [-0.41, 0.36]) to very high (e.g. Workspace 'maximum distance to front': ICC=0.95, 95% confidence interval [0.89, 0.97]). SRD% (absolute reliability) ranged from 5.9% (pROM 'shoulder extension') to 41'810.1% (RPM 'resistance against forearm pronation').

Conclusions: Relative and absolute test-retest reliability ranged widely within and between the assessments. For participants with reduced strength, it was difficult to perform the QoM and Circle assessments properly. Depending on the reliability results and experience obtained during the measurement sessions, we would recommend applying QoM, Circle, and Workspace also in clinical practice.

Trial registration: This study was registered prospectively at ClinicalTrials.gov (identifier: NCT02443857).

BACKGROUND

Disorders of the central nervous system might result in impaired motor function, which could negatively influence the participation in leisure activities of children and adolescents with cerebral palsy (CP) [1]. In children with CP, the prevalence of upper limb involvement is high and amounts to 83% [2]. Therefore, assessing upper limb functions is important when evaluating children undergoing paediatric neurorehabilitation.

Rehabilitation technologies are implemented more and more in upper limb therapy programmes. The main advantages are the high number of movement repetitions, the repeatability with which functions can be practiced and measured, and the goal-oriented training content [3]. As robot-assisted training is usually combined with exergames, motivation [4–6] and active engagement of the children can be increased playfully [7, 8]. Most upper limb systems have been designed for mild to moderately affected patients and do not contain drives to assist the patient with full support when moving. A certain level of upper limb function is a prerequisite to train with such devices. Devices matching these criteria include, for example, the Armeo®Spring (Hocoma, Switzerland [7, 9, 10]), which can support the weight of the upper and lower arm by two springs or the Rapael (Neofect, South Korea [11, 12]), which uses gloves to interact with the exergames and provides no physical support at all.

Therefore, in a collaboration between our Rehabilitation Centre for Children and Adolescents of the University Children's Hospital Zurich in Switzerland and the Sensorimotor Systems Lab of the Federal Institute of Technology in Zurich, Switzerland (SMS Lab), we developed a new robot which was designed especially for children with more severely affected upper limb function [13]. This Children Arm Rehabilitation Mechatronic Interface (ChARMin) robotic device is an exoskeleton with drives to support shoulder, elbow, forearm, and wrist movements. It also provides virtual reality games to induce engagement and motivation and it has seven assessments to quantify various upper limb functions in a standardised manner. Four assessments include functions, which are usually covered by conventional therapeutic assessments: 1) active range of motion (aROM) 2) passive range of motion (pROM), 3) isometric strength (Strength), and 4) resistance to passive movement (RPM) to quantify spasticity. The other assessments measure more complex movement functions, which are difficult to assess using routine clinical assessments: 5) quality of goal-directed movements (QoM), 6) dynamic tracking ability of the hand during a circle following task (Circle), and 7) workspace, where we evaluate the maximally reached distances in six movement directions.

A recently published systematic review on kinematic assessments of upper limb movements after stroke concluded that evaluating clinimetric properties of assessments is urgently needed [14]. The same applies to paediatric

neurorehabilitation, where we need to know the reliability of an assessment to be sure whether a measured change in a patient from pre- to post-treatment exceeds the 'normal' variability of the assessment.

Therefore, in this psychometric study, we aimed to establish the relative and absolute reliability of the ChARMin assessments. Additionally, we discuss the advantages and disadvantages of these assessments compared to routine clinical ratings based on the experiences obtained during this study.

METHODS

Participants

Participants were recruited among the in- and outpatients of the Swiss Children's Rehab, University Children's Hospital Zurich, Switzerland.

Inclusion criteria were: a) age 5 to 18 years, b) congenital or acquired brain lesion affecting upper limb function, c) ability to understand and follow test instructions, d) ability to sit upright for at least 60 minutes without lateral trunk support, e) Manual Ability Classification System (MACS) level I-IV (level I: handles objects easily and successfully, level II: handles most objects with somewhat reduced quality and/or speed of achievement, level III: handles objects with difficulty: needs help to prepare and/or modify activities, level IV: handles a limited selection of easily managed objects in adapted situations) [15].

Exclusion criteria were: a) severe obesity (i.e. upper limb too large for the robot's cuffs), b) fixed upper limb joint contractures, c) severe spasticity with Modified Ashworth Scale (MAS) ≥ 4 [16], d) unstable bones or joints, fractures or osteoporosis/ osteopenia, (sub-)luxations, e) upper limb surgery or botulinum toxin injections during the preceding 6 months, f) skin lesions, g) implanted devices (e.g. pacemakers, defibrillators), h) absence of compliance and inability to signal pain or discomfort, i) severe cognitive deficits, j) severe visual impairments.

Participants and legal representatives agreed verbally. Legal representatives and participants aged 14 years and older signed written informed consent. The study was approved by the Ethics Committee Zurich (BASEC-Nr. PB_2016-02450) and the Swiss Agency for Therapeutic Products (Swissmedic reference number: 2015-MD-0009).

Materials and procedure

The ChARMin robot and the assessments

ChARMin is an exoskeleton robot for training upper limb functions [13]. It is attached at the patient's upper arm and forearm with two hook-and-loop-fastener cuffs. The design is modular. A smaller or a larger distal module can be used with a child or adolescent, respectively, and optimally adjusted to the individual anthropometrics of each patient (Figure 3.1). ChARMin operates with three support modes (non-

supported, assist-as-needed, fully-guided). The support by the robot can be set between 0% and 100%, which enables training of children with a wide range of impairment severities. ChARMin has six actuated degrees of freedom, which can be moved independently: horizontal ab-/adduction, flexion/extension, and inner-/outer rotation of the shoulder; flexion/extension of the elbow; pro-/supination of the forearm; and flexion/extension of the wrist. ChARMin is interfaced with a computer screen on which different games and assessments can be visualised. ChARMin provides seven assessments to evaluate a wide spectrum of upper limb functions. Six assessment interfaces (aROM and pROM have the same interface) are displayed on Figure 3.2.

a) 'aROM' records the active range of motion of the shoulder, elbow, and wrist. The child is instructed to move the arm joints actively in the movement directions indicated on the screen.

b) 'pROM' records the passive range of motion of the shoulder, elbow, and wrist. The therapist moves the arm joints of the child in the movement directions indicated on the screen.

c) 'Strength' records the maximum isometric force of muscle groups. The exoskeleton remains in a static position, while the child applies maximal force in the joint direction indicated on the screen.

d) 'RPM' measures the resistance against movements of different speeds. After the instructing the children to keep their arm relaxed, the robot moves the child's arm in each joint direction at a speed of 10°/s and 60°/s.

e) 'QoM' measures the quality of goal-directed movements such as the precision or the fluency of the movement, but also the time needed to perform the tasks and the reaction time. The child is asked to move the hand from the centre point on the screen to eight different target points appearing one after the other, radially around the centre.

f) During 'Circle', the child is asked to position the red ball as precisely as possible on the green ball which is displayed on the screen and follow its circular movement by moving the arm accordingly with the attached exoskeleton.

g) 'Workspace' captures the active workspace of the arm in 3D. The child is instructed to 'push' the walls, ceiling, and floor of a virtual room displayed on the screen as far away as possible, making the room as large as possible.

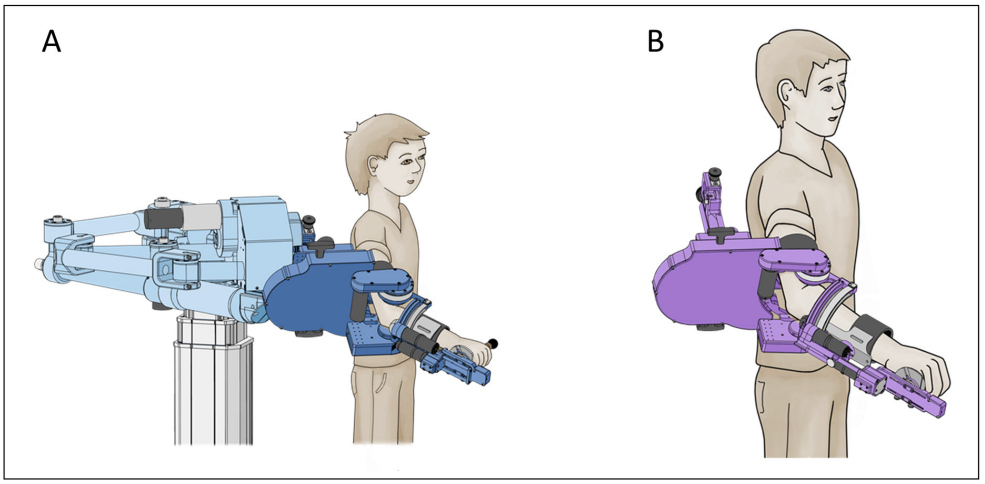


Fig. 3.1. Schematic picture of the ChARMin robot. A) Small distal module, B) large distal module.

Procedure

The participants attended two measurement sessions, scheduled three to seven days apart, each lasting one hour. The timespan was chosen to obtain stable yet independent measurements. During both sessions, a therapist guided each assessment verbally.

Measurement session 1: ChARMin was adjusted according to the participant's anthropometrics of the more affected arm. If both sides were similarly affected, the dominant side was chosen, as training this arm would be clinically more meaningful. After attaching the exoskeleton, the seven assessments were performed in random order, except for the pROM, which was always performed before aROM (since both are included in one ROM assessment) and the RPM. RPM is performed after pROM due to safety reasons, to ensure that the RPM is only moving the joints in the range obtained during pROM.

While the assessments aROM, pROM, Strength, RPM and Workspace, have parameter-specific instructions (e.g. aROM elbow flexion: 'bend your elbow as much as possible'), QoM and Circle are instructed in a more general matter (e.g. QoM: 'try to reach each appearing target') and the parameters that are recorded are not completely clear to the children. This could lead to different levels of performance influencing reliability. To achieve more stable data, QoM and Circle were repeated three times each and the average value of each parameter was used for analyses. All the assessments were, whenever possible, adjusted to the abilities of each child (e.g. the speed of the Circle assessment was reduced from default speed level 5 to speed level 3 if the participant was unable to follow the ball).

Measurement session 2: All the assessments were repeated in the same order and with the same settings of the exoskeleton and the assessments as during the first appointment.

The parameters we chose for analysis are listed in Table 3.1. For the RPM parameters, we calculated the differences of the resistance obtained during the slow and fast passive movement in line with the definition of spasticity (i.e. a velocity-dependent increase in muscle tone [17]) for the analysis. We adjusted the signs so that in all parameters a positive value means increase in tone and a negative value no increase in tone. Workspace measures the movement directions left and right side. We changed the parameters to “medial” and “lateral” direction since moving to the left, for example, for some children requires an adduction and for the others an abduction, depending on the assessed arm. Raw data of Strength and Workspace were measured as positive and negative numbers depending on the direction the force was applied and the spatial direction, respectively. However, for the analysis of these two assessments, absolute values were used.

The codes of the evaluation software used to calculate the parameters is available on the figshare database (doi: 10.6084/m9.figshare.9741221).

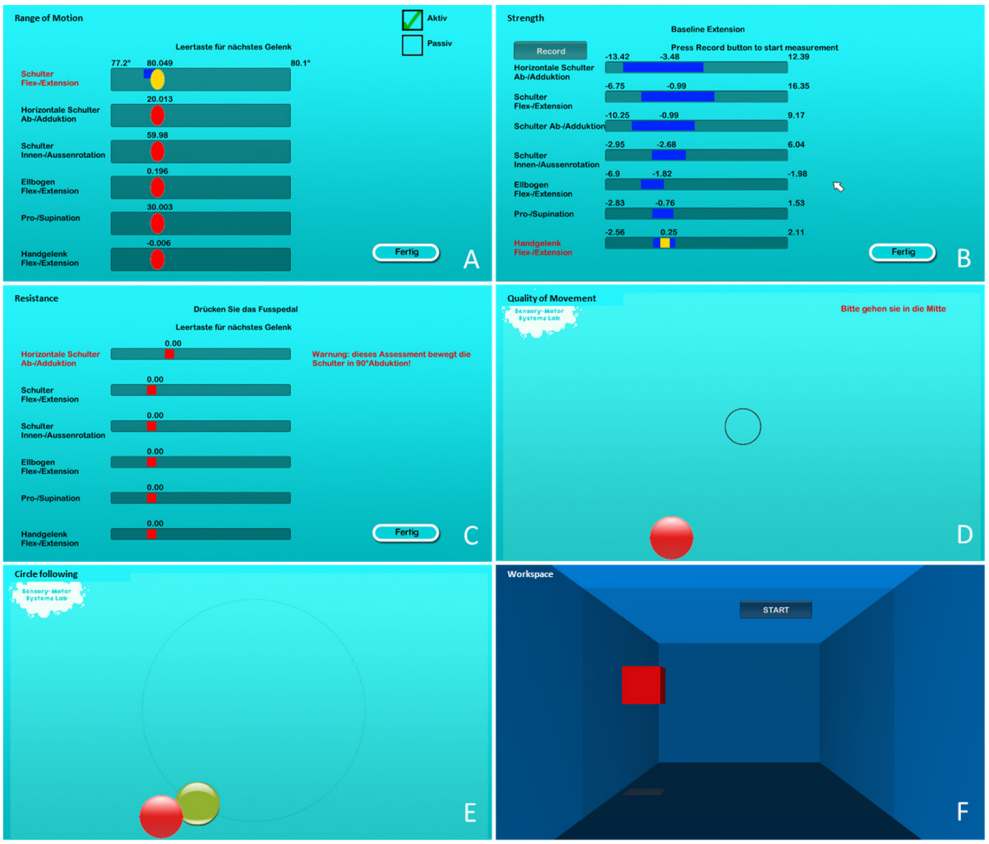


Fig. 3.2. Interfaces of the assessments. A) Active and passive Range of Motion. B) Isometric Strength. C) Resistance to Passive Movement. D) Quality of Movement: eight targets appearing radially around the centre point need to be reached. After each target, the participant has to return to the centre position. E) Circle following: the green ball moves in a circle and the participant is instructed to position the red ball as exactly as possible on the green ball throughout the circular movement. F) Workspace: the participant is instructed to make the virtual room on the screen as large as possible by pushing with the red block against each wall (in forward, backward, left and right direction, respectively), the ceiling upwards and the floor downwards. The block represents the position of the wrist and is steered by moving the arm in the according direction.

Table 3.1. Parameters of the ChARMin assessments

Assessment	Parameter [unit]	Description
Range of motion (active and passive)	Shoulder horizontal adduction [°]	Maximal range of motion of the shoulder in horizontal adduction
	Shoulder horizontal abduction [°]	Maximal range of motion of the shoulder in horizontal abduction
	Shoulder extension [°]	Maximal range of motion of the shoulder in extension
	Shoulder flexion [°]	Maximal range of motion of the shoulder in flexion
	Shoulder internal rotation [°]	Maximal range of motion of the shoulder in internal rotation
	Shoulder external rotation [°]	Maximal range of motion of the shoulder in external rotation
	Elbow extension [°]	Maximal range of motion of the elbow in extension
	Elbow flexion [°]	Maximal range of motion of the elbow in flexion
	Forearm pronation [°]	Maximal range of motion of the forearm in pronation
	Forearm supination [°]	Maximal range of motion of the forearm in supination
	Wrist flexion [°]	Maximal range of motion of the wrist in flexion
	Wrist extension [°]	Maximal range of motion of the wrist in extension
Strength	Shoulder horizontal abductors [Nm]	Isometric joint torques in the direction of shoulder horizontal abduction
	Shoulder horizontal adductors [Nm]	Isometric joint torques in the direction of shoulder horizontal adduction
	Shoulder extensors [Nm]	Isometric joint torques in the direction of shoulder extension
	Shoulder flexors [Nm]	Isometric joint torques in the direction of shoulder flexion
	Shoulder abductors [Nm]	Isometric joint torques in the direction of shoulder abduction
	Shoulder adductors [Nm]	Isometric joint torques in the direction of shoulder adduction
	Shoulder external rotators [Nm]	Isometric joint torques in the direction of shoulder external rotation
	Shoulder internal rotators [Nm]	Isometric joint torques in the direction of shoulder internal rotation
	Elbow extensors [Nm]	Isometric joint torques in the direction of elbow extension
	Elbow flexors [Nm]	Isometric joint torques in the direction of elbow flexion
	Forearm supinators [Nm]	Isometric joint torques in the direction of forearm supination
	Forearm pronators [Nm]	Isometric joint torques in the direction of forearm pronation

(Continued)

Table 3.1. Continued

Assessment	Parameter [unit]	Description
Strength *(Continued)*	Wrist extensors [Nm]	Isometric joint torques in the direction of wrist extension
	Wrist flexors [Nm]	Isometric joint torques in the direction of wrist flexion
	Hand/finger flexors [Nm]	Isometric joint torques in the direction of hand/finger flexion
Resistance to passive movement (at 60°/s minus resistance occurring at 10°/s)	Against shoulder horizontal abduction [Nm/rad]	Resistance occurring during shoulder horizontal abduction
	Against shoulder horizontal adduction [Nm/rad]	Resistance occurring during shoulder horizontal adduction
	Against shoulder flexion [Nm/rad]	Resistance occurring during shoulder flexion
	Against shoulder extension [Nm/rad]	Resistance occurring during shoulder extension
	Against shoulder external rotation [Nm/rad]	Resistance occurring during shoulder external rotation
	Against shoulder internal rotation [Nm/rad]	Resistance occurring during shoulder internal rotation
	Against elbow extension [Nm/rad]	Resistance occurring during elbow extension
	Against elbow flexion [Nm/rad]	Resistance occurring during elbow flexion
	Against forearm supination [Nm/rad]	Resistance occurring during forearm supination
	Against forearm pronation [Nm/rad]	Resistance occurring during forearm pronation
	Against wrist extension [Nm/rad]	Resistance occurring during wrist extension
	Against wrist flexion [Nm/rad]	Resistance occurring during wrist flexion
Quality of movement (the mean over all movements)	Mean distance-to-path-ratio [unitless]	Length of the trajectory from start/target to the target/start divided by the direct distance between the start and the target position
	Mean standard deviation [m]	Standard deviation of the end-effector position is calculated for the time when the patient's hand is on the target position
	Mean number of peaks [n speed peaks/distance]	Number of speed peaks normalised to the trajectory distance
	Mean absolute number of peaks [n speed peaks]	Number of speed peaks
	Mean time [ms]	Difference between the two timestamps when the patient reaches the target and when the patient left the starting position

(Continued)

Table 3.1. Continued

Assessment	Parameter [unit]	Description
Quality of movement (the mean over all movements) *(Continued)*	Mean reaction time [ms]	Time between the timestamp when the target is shown and the time when the robot end effector leaves the start/target position
Circle (mean of all the trials)	Mean summed difference [m]	Summed difference between the current position and the reference circle
	Mean percentage in front [%]	Percentage of the time that the participant was in front of the reference circle
	Mean ellipse ratio [unitless]	Ratio between the minimal and the maximal radius of the least-squares fitted ellipse
Workspace	Maximum distance lateral [m]	Maximal displacement in direction away from the body, (lateral direction)
	Maximum distance down [m]	Maximal displacement in direction of the bottom wall
	Maximum distance medial [m]	Maximal displacement in direction to the body (medial direction)
	Maximum distance up [m]	Maximal displacement in direction of the top wall
	Maximum distance chest [m]	Maximal displacement in direction towards the chest
	Maximum distance front [m]	Maximal displacement in direction of the front wall
	Volume [m^3]	Cubic volume of the arm reachable workspace

Parameters obtained from the ChARMin assessments, their unit and the explanation of each parameter. Abbreviations: °=degrees, Nm=Newton-metre, Nm/rad=Newton-metre per radian, m=metre, n=number, ms=milliseconds, %=percentage, m^3=cubic metre.

Data processing and statistical analysis

The data, which was measured during the assessments, required further offline processing to extract the assessment parameters. We wrote the customised algorithms in MATLAB (R2014a, and R2017a, The MathWorks, Inc.).

The statistical analyses were performed with IBM SPSS Statistics 24. We tested data for normal distribution (Shapiro-Wilk-Test). Test-retest reliability was evaluated following a 3-layered approach [18]:

1. **Testing for systematic error:** We applied a paired T-test or Wilcoxon signed-rank test, depending on the data distribution, to test for systematic error between session 1 and 2.
2. **Relative reliability:** We used a two-way mixed model Intraclass Correlation Coefficient (ICC), type absolute agreement [19]. For QoM and Circle we selected the average data option in the model, for the other assessments, we selected single data. Even if data were not normally distributed, ANOVAs are relatively robust to violations of this assumption [20].
3. **Absolute reliability:** The absolute Smallest Real Difference (|SRD|) was calculated based on variance values obtained from the ANOVA table:

$$|SRD| = SEM \times 1.96 \times \sqrt{2} \quad [18]$$

$$SEM = \sqrt{(s^2_{residual})} \quad [21]$$

We calculated |SRD%|, the percentage of the |SRD| of the Grand Mean (GM; i.e. average of the first and second measurement):

$$SRD\% = \frac{|SRD|}{GM} \times 100$$

RESULTS

Thirty participants, 9 females and 21 males, were included. Their mean age ± standard deviation (SD) was 12.5 ± 3.3 years. Their height amounted to 139.0 ± 42.3 cm and their weight to 46.9 ± 20.3 kg (mean ± SD). Diagnoses were congenital (n=15) or acquired brain injury (n=14) or both (n=1). The severity levels were MACS level I: n=8, MACS level II n=12, MACS level III: n=8, MACS level IV: n=1, for one participant the MACS level was not available. Twenty-three participants were inpatients, and seven were outpatients. With the exception of three participants, the measurement sessions were performed on the same half of day.

The original data table is available on the figshare database (doi: 10.6084/m9.figshare.9741221).

Additional figure files show scatterplots indicating the relationships between the first and second measurements and the distribution of the data of each assessment and parameter (see Additional files 3.1 to 3.6).

1. **Systematic error:** Wilcoxon signed-rank tests confirmed that there were no systematic errors between the first and second assessments of aROM, Strength, Circle, and Workspace (Table 3.2). Four of six parameters of the QoM assessment (i.e. 'mean distance-to-path-ratio', 'mean standard deviation', 'mean absolute number of peaks', 'and mean time') were significantly different between the sessions and indicated that QoM was performed more fluently, with a more direct and faster movement during the second session. The parameters 'number of peaks' (normalised to distance) and 'reaction time' did not differ between the sessions. For the other assessments, only one parameter was significantly different between the two sessions (Table 3.2).
2. **Relative reliability:** Overall, ICCs ranged from little to very high (Table 3.2). The highest ICC was obtained in Workspace 'maximum distance to front' (ICC = 0.95, 95% confidence interval (95%CI) [-0.41, 0.36]), the lowest ICC in RPM 'resistance against shoulder external rotation' (ICC = - 0.03, 95%CI [-0.41, 0.36]). Fig 3.3 shows the relationship between the measurements and the distribution of the data of these two parameters.
3. **Absolute reliability:** SRDs for assessment parameters ranged widely within but also between the assessments (Table 3.2, Fig 3.4). The lowest |SRD%| was found for the pROM parameter 'shoulder extension' (5.9%) and highest for the RPM parameter 'resistance against forearm pronation' (41'810.1%).

Table 3.2. Systematic error, relative and absolute reliability of the parameters of the ChARMin assessments

Assessment	Parameter [unit]	n datasets in analysis	Systematic error		Relative reliability				Absolute reliability	
			WSR test		ICC	95%CI		p-value	SRD	|SRD%|
			Median Diff. (M1-M2)	p-value		LB	UB			
Active range of motion	Shoulder horizontal adduction [°]	23	-0.23	0.36	0.19	-0.25	0.56	0.19	2.88	35.07
	Shoulder horizontal abduction [°]	23	0.34	0.76	0.49	0.09	0.75	0.01	23.84	27.18
	Shoulder extension [°]	27	0.17	0.27	0.70	0.44	0.85	<0.001	3.18	6.07
	Shoulder flexion [°]	27	0.49	0.07	0.88	0.75	0.94	<0.001	17.72	15.07
	Shoulder internal rotation [°]	23	-0.22	0.88	0.54	0.17	0.78	<0.01	13.14	46.80
	Shoulder external rotation [°]	23	0.43	0.07	0.90	0.79	0.96	<0.001	15.73	26.45
	Elbow extension [°]	27	8.60	0.12	0.57	0.26	0.78	<0.001	35.00	213.49
	Elbow flexion [°]	27	-0.13	0.68	0.82	0.64	0.91	<0.001	8.50	7.46
	Forearm pronation [°]	25	-10.12	0.23	0.58	0.26	0.79	<0.001	34.23	189.10
	Forearm supination [°]	25	37.77	0.06	0.90	0.77	0.95	<0.001	33.59	88.96
	Wrist flexion [°]	28	-0.69	0.37	0.66	0.39	0.83	<0.001	19.32	33.11
	Wrist extension [°]	28	-11.09	0.39	0.86	0.72	0.93	<0.001	25.72	68.82
Passive range of motion	Shoulder horizontal adduction [°]	28	-0.04	0.94	0.55	0.23	0.76	<0.01	4.56	67.50
	Shoulder horizontal abduction [°]	28	-0.64	0.12	0.71	0.47	0.85	<0.001	17.15	20.06
	Shoulder extension [°]	29	0.06	0.54	0.74	0.52	0.87	<0.001	3.10	5.93
	Shoulder flexion [°]	29	-0.12	0.30	0.59	0.28	0.78	<0.001	19.39	15.80
	Shoulder internal rotation [°]	27	-0.06	0.95	0.01	-0.38	0.39	0.48	4.29	14.82
	Shoulder external rotation [°]	27	0.04	0.59	0.17	-0.22	0.51	0.20	7.69	11.70
	Elbow extension [°]	29	0.14	0.13	0.71	0.45	0.86	<0.001	13.79	136.28
	Elbow flexion [°]	29	-0.11	0.18	0.14	-0.23	0.48	0.23	9.98	8.68
	Forearm pronation [°]	26	10.62	0.10	0.54	0.21	0.76	<0.01	40.29	60.71
	Forearm supination [°]	26	0.22	0.44	0.43	0.06	0.70	0.01	40.28	53.98
	Wrist flexion [°]	29	0.20	0.21	0.54	0.23	0.75	<0.01	15.45	27.23
	Wrist extension [°]	29	-10.29	0.03	0.65	0.38	0.82	<0.001	23.55	48.27

(Continued)

Table 3.2. Continued

Assessment	Parameter [unit]	n datasets in analysis	Systematic error		Relative reliability				Absolute reliability	
			WSR test		ICC	95%CI		p-value	SRD	|SRD%|
			Median Diff. (M1-M2)	p-value		LB	UB			
Strength	Shoulder horizontal abductors [Nm]	28	-3.74	0.30	0.92	0.84	0.96	<0.001	6.67	58.33
	Shoulder horizontal adductors [Nm]	26	1.58	0.09	0.87	0.71	0.94	<0.001	8.58	62.12
	Shoulder extensors [Nm]	27	2.54	0.96	0.68	0.42	0.84	<0.001	16.15	178.76
	Shoulder flexors [Nm]	27	-3.12	0.18	0.83	0.67	0.92	<0.001	16.15	68.08
	Shoulder abductors [Nm]	26	-4.10	0.66	0.34	-0.05	0.63	0.0	20.94	239.48
	Shoulder adductors [Nm]	27	-2.26	0.16	0.64	0.35	0.81	<0.001	17.44	175.02
	Shoulder external rotators [Nm]	27	0.42	0.55	0.90	0.80	0.95	<0.001	4.72	56.20
	Shoulder internal rotators [Nm]	27	-0.37	0.09	0.69	0.43	0.84	<0.001	7.88	119.49
	Elbow extensors [Nm]	27	-1.52	0.12	0.64	0.36	0.82	<0.001	9.48	120.02
	Elbow flexors [Nm]	27	-0.62	0.90	0.89	0.77	0.95	<0.001	7.13	63.01
	Forearm supinators [Nm]	24	0.15	0.49	0.81	0.61	0.91	<0.001	1.18	134.90
	Forearm pronators [Nm]	24	-0.33	0.95	0.61	0.27	0.81	<0.001	2.11	194.72
	Wrist extensors [Nm]	28	-0.17	0.73	0.64	0.35	0.82	<0.001	2.16	156.37
	Wrist flexors [Nm]	28	0.15	0.96	0.60	0.30	0.79	<0.001	2.96	144.37
	Hand flexors [N]	28	0.00	0.40	0.70	0.45	0.85	<0.001	1.07	132.98
Resistance to passive movement	Against horizontal shoulder abduction [Nm/rad]	26	-0.13	0.69	0.19	-0.22	0.54	0.18	4.02	298.54
	Against horizontal shoulder adduction [Nm/rad]	26	0.16	0.01	0.76	0.51	0.89	<0.001	2.87	526.79
	Against shoulder flexion [Nm/rad]	26	-2.04	0.29	0.70	0.43	0.85	<0.001	14.79	672.21
	Against shoulder extension [Nm/rad]	26	1.12	0.18	0.25	-0.15	0.58	0.10	10.37	1533.00
	Against shoulder external rotation [Nm/ rad]	26	0.26	0.59	-0.03*	-0.41	0.36	0.55	3.05	2570.99
	Against shoulder internal rotation [Nm/rad]	26	0.00	0.85	0.72	0.47	0.86	<0.001	0.87	375.90

(Continued)

Table 3.2. Continued

Assessment	Parameter [unit]	n datasets in analysis	Systematic error		Relative reliability				Absolute reliability	
			WSR test		ICC	95%CI		p-value	SRD	|SRD%|
			Median Diff. (M1-M2)	p-value		LB	UB			
Resistance to passive movement *(Continued)*	Against elbow extension [Nm/rad]	26	0.05	0.14	0.52	0.18	0.75	<0.01	3.48	164.34
	Against elbow flexion [Nm/rad]	26	-0.08	0.42	0.17	-0.23	0.52	0.20	2.24	1165.18
	Against forearm supination [Nm/rad]	23	-0.02	0.47	0.70	0.41	0.86	<0.001	1.15	2634.61
	Against forearm pronation [Nm/rad]	23	0.07	1.00	0.43	0.01	0.71	0.02	0.77	41810.05
	Against wrist extension [Nm/rad]	26	0.04	0.77	0.22	-0.19	0.56	0.15	0.81	20839.08
	Against wrist flexion [Nm/rad]	26	-0.02	0.29	0.44	0.09	0.70	0.01	0.56	1894.62
Quality of movement	Mean distance-to-path-ratio [unitless]	23	0.12	<0.01	0.85	0.45	0.95	<0.001	0.64	35.58
	Mean standard deviation [m]	23	0.00	0.03	0.86	0.64	0.94	<0.001	0.01	61.73
	Mean number of peaks [n speed peaks/distance]	23	1.05	0.93	0.94	0.86	0.98	<0.001	4.01	17.71
	Mean absolute number of peaks [n speed peaks]	23	0.96	<0.01	0.79	0.32	0.92	<0.001	2.15	33.15
	Mean time [ms]	23	1473.01	<0.001	0.87	0.40	0.96	<0.001	2742.99	51.21
	Mean reaction time [ms]	23	259.43	0.10	0.69	0.20	0.88	<0.001	1032.25	81.26
Circle following	Mean summed difference [m]	14	0.01	0.70	0.42	-0.92	0.82	0.18	1.81	218.06
	Mean percentage in front [%]	14	-0.23	0.55	0.63	-0.12	0.88	0.04	39.88	75.54
	Mean ellipse ratio [unitless]	14	-0.01	0.59	0.73	0.18	0.91	0.01	0.29	37.27
Workspace	Maximum distance lateral [m]	29	-0.01	0.61	0.84	0.69	0.92	<0.001	0.11	29.22
	Maximum distance down [m]	29	0.00	0.27	0.71	0.46	0.85	<0.001	0.10	34.82
	Maximum distance medial [m]	29	0.00	0.30	0.74	0.53	0.87	<0.001	0.07	32.86
	Maximum distance up [m]	29	-0.04	0.27	0.92	0.84	0.96	<0.001	0.12	45.37
	Maximum distance to chest [m]	29	0.00	0.35	0.29	-0.07	0.59	0.06	0.13	58.58
	Maximum distance to front [m]	29	0.00	0.75	0.95	0.89	0.97	<0.001	0.04	8.58
	Volume [m^3]	29	-0.01	0.09	0.91	0.82	0.96	<0.001	0.06	61.50

(Continued)

Table 3.2. Continued

Results of the test for systematic error. Relative reliability and absolute reliability. Displayed are the analysed parameters of all assessments. Intraclass Correlation Coefficients of Quality of Movement and Circle assessment are based on average measures. Intraclass correlation coefficients of active and passive Range Of Motion, Strength, Resistance to Passive Movement and Workspace are based on single measures.

Abbreviations: WSR test = Wilcoxon signed-rank test; Diff. = difference; M1 = median of the first measurement; M2 = median of the second measurement; ICC= intraclass correlation coefficient; 95%CI = 95% confidence interval; LB = lower bound; UB = upper bound; SRD = smallest real difference; SRD% = smallest real difference/grand mean x 100, °=degrees, Nm=Newton-metre, Nm/rad=Newton-metre per radian, m=metre, n=number, ms=milliseconds, %=percentage, m^3=cubic metre.

*A negative ICC is referred to as a bad or unfortunate estimate, possibly occurring by chance, especially with a small sample size [22].

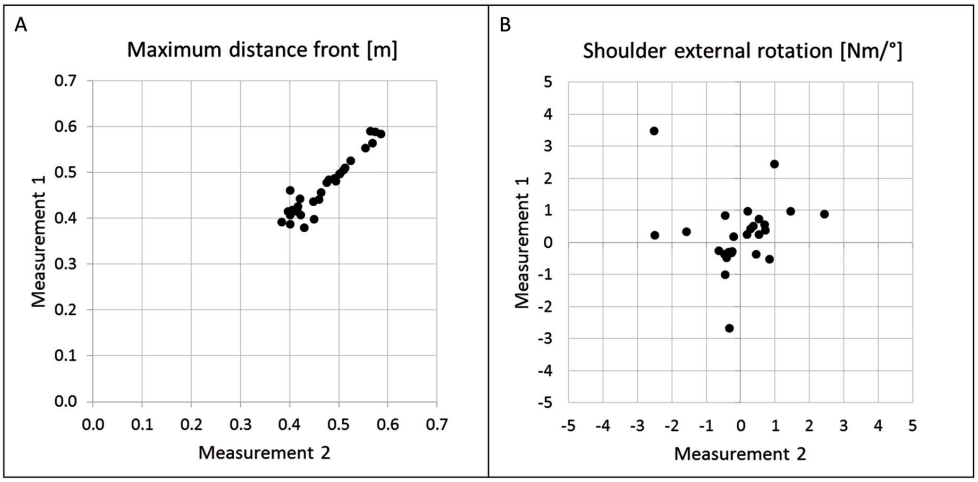


Fig. 3.3. Data distribution of the parameters with the highest and the lowest intraclass correlation coefficients (ICC). A) The parameter 'maximum distance to front' of the Workspace assessment which showed the highest ICC (= 0.95). B) The parameter 'resistance against shoulder external rotation' of the Resistance to Passive Movement assessment (RPM) which showed the lowest ICC (=- 0.03).

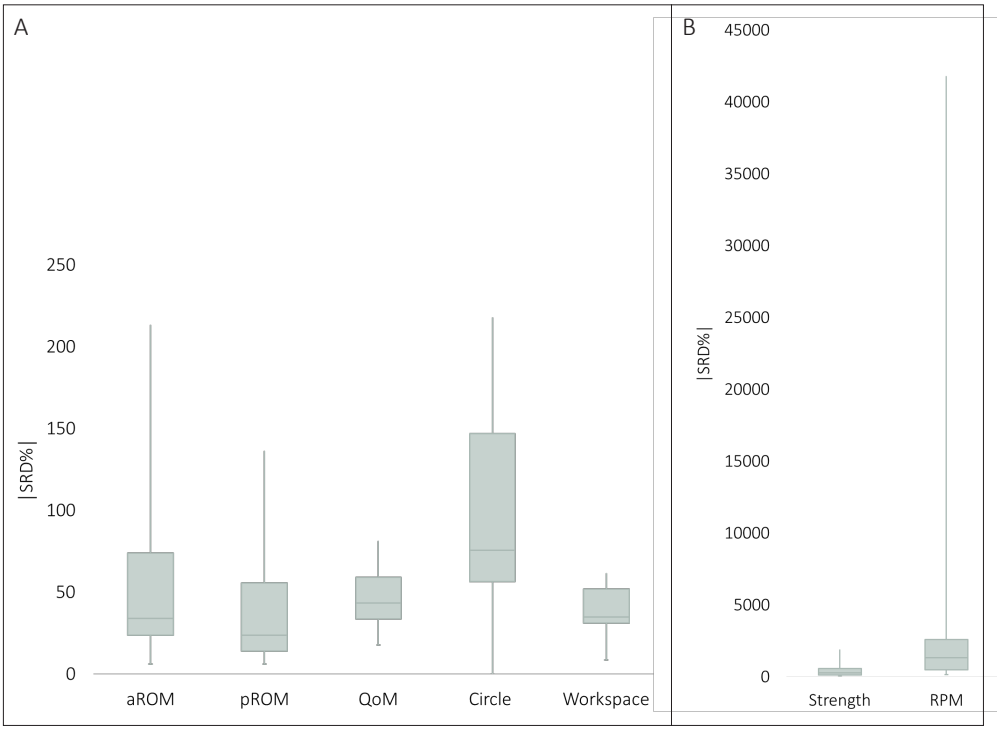


Fig. 3.4. Measurement errors. The box-plots represent the distribution of the smallest real differences as a ratio of the grand means of all the parameters of each assessment. A) Active Range of Motion (aROM), passive Range of Motion (pROM), Strength, Quality of Movement (QoM), Circle, and Workspace assessments. B) Resistance against passive movement (RPM).

DISCUSSION

We evaluated the relative and absolute test-retest reliability of seven assessments integrated in the ChARMin robot evaluating upper limb functions in children undergoing neurorehabilitation. Analysis of relative and absolute reliability showed large differences between the parameters and assessments. In the following paragraphs, we will discuss our results in relation to previous study findings and our clinical experience.

aROM and pROM

The conventional method to measure ROM is using a manual goniometer. The intra-rater reliability of goniometry of active movements was evaluated in children after forearm fractures [23]. ICCs ranged from 0.73 to 0.97 for pronation and from 0.80 to 0.97 for supination in different age groups [23]. While the result for supination from our study is in line with these results, pronation was less reliable in our study. There are many factors potentially affecting reliability such as the patient population (impairments, cognitive abilities influencing the test performance),

limitations of the robotic device, the level of variability between the participants or the number of repetitions of a measurement. Colaris et al. compared the means of three measurements [23] which might have led to more stable numbers since the influence of single outliers might be reduced, while our results were based on one ROM measurement. In the same study, the SRD ranged from 4 to 9 degrees for pronation and 5 to 9 degrees for supination [23]. These results indicate a better absolute reliability compared to our study. Even though both studies have used the same calculations for the SRD, the SEM on which these calculations are based, might have been calculated differently influencing the SRD value.

In a test-retest reliability study with children with CP, the ICC for pROM elbow extension was 0.94, 95%CI [0.86, 0.97], for wrist supination 0.81, 95%CI [0.61, 0.91], and for wrist extension 0.88, 95%CI [0.74 – 0.95] [24]. In a robotic pROM assessment in healthy adults, ICCs for wrist flexion were 0.97-0.98 and for wrist extension 0.87-0.95, while the SRD% varied between 9.9%-19.6% for both movement directions [25]. Reliability of our data is considerably lower. An explanation could be that in some movement directions (shoulder horizontal adduction, shoulder extension and flexion, shoulder internal and external rotation and elbow flexion) most of the participants reached the mechanically maximally possible pROM. Variability between the participants was therefore small which is partly reflected in rather small ICCs and SRDs (Table 3.2).

Strength

Muscle strength is conventionally measured with the manual muscle test [26]. In children with CP, this assessment showed a test-retest reliability with ICCs from 0.69, 95%CI [0.40, 0.85] for the shoulder abductors to 0.98, 95%CI [0.95, 0.99] for forearm supinators [24]. Our results indicate lower reliability. While compared to the other ChARMin assessments, the relative reliability of Strength was rather good, the absolute reliability showed partly large measurement errors (Table 3.2).

Participants were instructed to keep still during a baseline measurement and then to push against the robot in the according direction. For some of them, this was difficult. Sometimes, they were not completely relaxed during the baseline period or they started too early pushing against the robot, which influenced the baseline value, and therewith the strength value (as we subtracted the baseline value). If too much force is applied to the robot, the software stops and the measurement needs to be repeated. This was a problem with participants with a high level of strength. When the assessment had to be repeated, participants often reduced their force in order to prevent the robot from stopping which distorted our results. Since, regarding strength, we had a heterogeneous study sample, the widespread distribution of strength levels might have influenced the ICCs positively, while the errors occurring during the baseline measurements might reflect the high measurement errors. This divergence of relative and absolute reliability can be seen in the majority of the

Strength parameters (Table 3.2, Additional file 3.2). One example is the parameter 'shoulder horizontal abductors' which yielded the highest ICC of all the Strength parameters, yet, a |SRD%| of 58.3%.

Measuring strength with ChARMin would provide objective data, which is not the case when using a manual muscle test. However, as the Strength assessment is now, there are too many uncertainties in order to use the data to determine outcomes. We would recommend making the application of the strength assessment easier to follow. i.e. to measure the baseline after the isometric force measurement so that it is easier for the children to relax.

RPM

The MAS is commonly used to test spasticity. Test-retest reliability was evaluated in children with CP and showed ICCs between 0.70, 95%CI [0.42, 0.86] for shoulder adductors to 0.85,95%CI [0.69, 0.93] for elbow flexors [24].

Spasticity is defined as an increase in velocity-dependent stretch reflex [27]. It is recommended to stretch the muscle to be tested from one end position to the other within one second [16]. To test the elbow flexors, for example, this means a speed of about 180°/s if the patient has no joint contractures. The RPM assessment provided by ChARMin moves with 60°/s at fast speed. This limitation in movement velocity is mainly implemented due to safety reasons. However, it remains unclear whether an increase in resistance can be provoked with this movement speed. The validity of the RPM assessment needs to be evaluated. Also, in most of the participants ChARMin did not move through the full range of motion since it is restricted due to safety reasons. Hence, end-range movements are not tested and it is not possible to detect a catch and release phenomenon at the end of a movement as it is done with the MAS [16], which is quite commonly used to test for spasticity. The advantage of RPM is that the data are interval-scaled and not ordinal-scaled like the MAS and, therefore, can be analysed parametrically. However, compared to the RPM parameters, test-retest reliability of the MAS showed much better reliability in children with CP [24]. Another study concluded the Ashworth Scale as not valid and reliable enough to be used to measure spasticity after evaluating it in patients with upper motor neuron syndrome [28]. While a spasticity measure with a robot would increase objectivity immensely, during a "hands-on" assessment, qualitative properties of the resistance might be noticed more easily than with the robot-assisted assessment. Furthermore, it is very difficult for the children to relax their arms completely during RPM. Their arm is attached to a moving exoskeleton. This is unfamiliar and some children experienced it as frightening. Therefore, it remains unclear whether we actually assessed increased muscle tone during the faster movement, or active resistance against the movement, which might have increased during the faster perhaps more uncomfortable movement in the robot or whether the resistance of the passive structures (e.g. ligaments) or shortened muscle structures are measured. The

differentiation between spasticity and passive structures as origin of the resistance is not possible. Muscles and passive soft tissues are viscoelastic, meaning that the resistance due to passive stretching is velocity dependent and increases with the velocity of a stretch, as spasticity does [29]. Therefore, it will always be difficult to differentiate spasticity from resistance of passive tissue especially when the former is of low severity.

Due to the low reliability and the ambiguities occurring during the application, we would not recommend to use RPM in clinical practice.

QoM

Measuring the quality of movement is an issue in a clinical setting. It starts with the question about what aspects of movement reflect movement quality. During the QoM assessment, many parameters are recorded which might reflect an aspect of movement quality. In a qualitative study, physiotherapists were interviewed about the phenomenon “movement quality” [30]. Among others, “the movement characteristic of path and form in movement” and “the movement characteristics of flow, elasticity, and rhythm” [30]. p. 17, were mentioned during the interviews. Based on that qualitative study, the parameters we chose for the analysis seem to be appropriate to quantify quality of movement.

Parameters of the QoM assessment measured with the precursor device of ChARMin in healthy adults have already shown a tendency to systematic error in intra-rater reliability [31]. It seems that in this study participants performed the QoM with a more fluent, more direct, and faster movement at session 2. Apparently, participants become more familiar with performing the test, which is reflected in improved performance of four parameters. However, as we did not correct for multiple comparisons false positive results might have occurred. When dividing the p-value of 0.05 with 67 (number of comparisons) the corrected p-valued would be 0.0007. Appropriately, only the T-test of the parameter ‘mean time’ would still be significant. Still, if the QoM assessment would be used to measure change in movement quality due to an intervention, this familiarization aspect should be taken into account. As the parameter number of speed peaks normalised to the actual path did not show such a familiarization, our results suggest that this parameter could be reliable enough to serve as an outcome parameter.

A similar point-to-point reaching task was evaluated with healthy adults with the MIT-Manus in the horizontal plane [32]. No systematic error was found in any of the variables and the authors concluded that all parameters are reliable [32]. However, they evaluated data from six measurement sessions on three days (two sessions per day) over two to three weeks [32]. Compared to children with neurological impairments, healthy adults are not or less affected by day-to-day variability due to spasticity, fatigue or motivation and differences in arm strength influence the performance less, because the movements were performed in the horizontal plane.

Therefore, these results are not directly comparable with our study.
The main difficulty when performing the QoM assessment was the lack of motivation. Reaching eight targets and repeating the procedure three times is not very exciting and children became quickly bored. Furthermore, while it was difficult reaching the upper targets for children with limited strength (Figure 3.5), adding support from ChARMin made it more difficult to reach the lower targets as the participants than had to push against the device downwards. We would recommend to make this assessment more suitable and interesting for children. However, as QoM parameters reflecting movement quality are difficult to obtain with current routinely applied clinical assessments, we consider this assessment of great interest when assessing upper limb function in paediatric neurorehabilitation. We recommend to improve its application and subsequently use QoM in clinical practice.

Circle
The Circle assessment enables as well to measure more qualitative movement parameters. The MIT-Manus also provides a similar reliable assessment [32], however, also here, the differences in study population and procedures make it difficult to compare the results.

Similar to the QoM, it was difficult to reach the full circle for children with difficulties moving the arm against gravity (Fig 5). Unfortunately, the processing software revealed a bug, which lead to the exclusion of datasets. The interpretation of the results obtained from 14 datasets is limited. Generally, and also similarly to the QoM, the Circle assessments would be able to measure clinically relevant parameters reflecting more qualitative aspects of movement. We recommend to use Circle in clinical practice with the exception of the parameter 'mean summed difference' due to its low reliability.

Workspace
This assessment provides important functional parameters that quantify reaching distance, which is an assessment of interest, because it can be linked well to various ADL.

Compared to the parameter 'maximal distance to front', the parameter 'maximum distance to chest' shows a very low reliability. An explanation for the difference in relative reliability could be that the movement direction to the chest has a relatively fixed endpoint (reaching the chest), while the endpoint of the movement in the front direction is more variable and differing between the participants leading to a high ICC.

The workspace volume has been measured and its reliability has been evaluated with the ARMEO Spring device in adult patients with neurological deficits in the upper limbs and healthy subjects [33]. Intra-rater reliability showed ICCs from 0.75

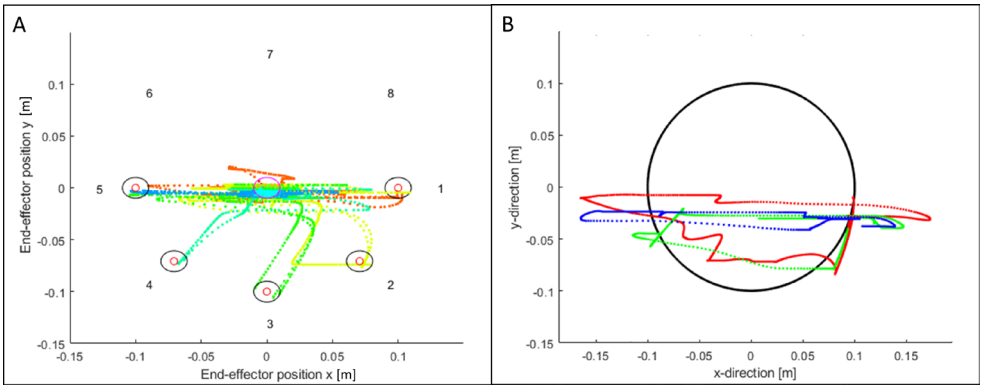


Fig. 3.5. Examples of trajectories of Quality of Movement and Circle assessments. The trajectories were obtained from data of an adolescent participant with acquired hemiparesis and a MACS level III who had difficulties in moving the arm upwards against gravity. A) Quality of Movement assessment: Paths for the movements from the targets to the centre point. Upper targets were not reached. B) Circle assessment: Paths of the three rounds of tracking the ball moving in a circle. The upper part of the circle was not reached. Red line = round one, green line = round two, blue line = round three. The excursion of the movement becomes smaller with each round.

to 0.86 for different seating positions (healthy adult subjects) [33]. These results are lower than the ICC obtained for the parameter volume in this study. The sample included in our study is probably more heterogeneous compared to the healthy subjects of the ARMEO Spring study. A wide variability between the participants might lead to a higher ICC. However, this should have resulted in higher ICC values in all parameters, which is not the case. In eight adults with neurological upper limb disorders (sitting in the chair used at the current phase of rehabilitation), the ICC was 0.86 [33]. Since they assessed both upper limbs of each participant, they included 14 datasets, which might not be considered completely independent, while in our study, 29 independent Workspace datasets were analysed. This might be another reason for the differences in ICCs.

One limitation of this assessment was that some participants were not able to reach the centre position by themselves, which is needed to start the assessment. The therapist had to support the child to reach the starting position, but this does not influence the data.

There is no conventional clinical assessment providing this information in such an easily applicable way. As the workspace might be relevant to improve for some children and since the parameters can be reliably measured, we recommend to use this assessment also in clinical practice with the exception of the parameter ‘maximum distance front’ due to its low reliability.

Clinical implications

Major advantages of the ChARMin and other robot-assisted devices are the objectivity of the measurements and the application of assessments to measure upper limb functions, which are usually not covered by conventional assessments. However, according to our experience, it should carefully be pondered when the use of ChARMin is reasonable. As we stated before, we would recommend QoM, some parameters of Circle and most of the parameters of Workspace for the clinical use, as there are no conventional assessments measuring these parameters and also because they demonstrated sufficient reliability.

Since adapting the exoskeleton to the anthropometry of the child and creating a user profile is quite time-consuming, it makes more sense to conduct assessments with children who are already training with the ChARMin or with whom training is planned. Applying only the assessments is therefore not recommended. Furthermore, the application of devices in rehabilitation is considered a continuum, in terms of starting with a device with more passive features and switching to another device when functions have improved [9, 34]. ChARMin, for example, is suitable during a specific period and its use is replaced by another tool providing less support as soon as this is required to progress with the course of rehabilitation. Therewith, it would also make sense to use ChARMin assessments during this time-period combined with some conventional assessments that can be applied continuously during the whole rehabilitation process, such as the Melbourne Assessment 2 [35] or the Assisted Hand Assessment [36].

Preferably, future robotic assessments could be implemented in an exergame, allowing to record outcome parameters while playing. This procedure would have the advantage of keeping the children engaged and motivated during an assessment.

Limitations

We included participants with congenital or acquired brain injuries affecting upper limb function. However, in two cases, the diagnoses were not completely clear as both participants had several comorbidities. Both had a brain damage involvement but the actual reason for the impaired upper limb function could also have been influenced by one of these comorbidities. By including these participants, we increased the heterogeneity of the study sample but this also reflects the population of children undergoing neurorehabilitation.

Some technical problems (e.g. bugs in software), compliance issues of the participants, and the inability to perform some assessments led to missing data. ChARMin was developed for children with more severe impairments. However, our study showed that participants already require certain abilities to perform the assessments. ChARMin is a unique specimen and not commercially available so far. Despite this, our findings and experiences might be useful for manufacturers of similar devices when developing and implementing assessments for the use in clinical practice.

CONCLUSIONS

Relative and absolute test-retest reliability of seven assessments evaluating upper limb functions provided by the ChARMin is varying between and within the assessments. We do not recommend to apply aROM and pROM as ROM is limited mechanically due to safety, and RPM due to low reliability. Strength shows potential for clinical use if data evaluation and measuring procedure are adapted accordingly. We recommend using the QoM, Circle, and Workspace assessments in clinical practice in those patients who use the device during their therapy also because there are no alternatives among the conventional assessments evaluating the respective parameters. However, in more severely impaired children, for whom ChARMin was actually designed, some assessments proved to be (too) difficult.

ACKNOWLEDGEMENTS

We would like to thank all the children who participated in this study and their parents. Special thanks goes to the occupational therapists Jan Lieber and Annina Herzog who performed some of the measurements, to Caroline Bastiaenen for her support with the manuscript and to Robert Riener and the Sensory-Motor Systems Lab at the ETH with whom we developed and manufactured the ChARMin exoskeleton.

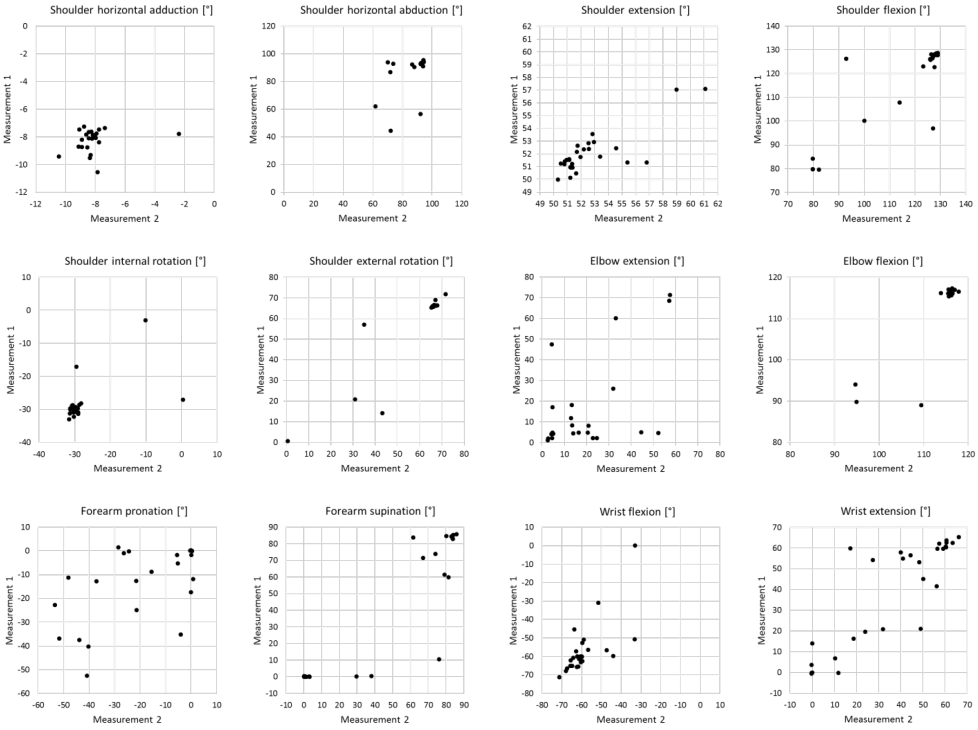
AUTHOR CONTRIBUTIONS

Conceptualization: Hubertus JA van Hedel, Judith V. Graser, Laura Prospero, Monica Liesch
Formal analysis: Judith V Graser
Funding acquisition: Hubertus JA van Hedel
Investigation: Judith V Graser, Laura Propsero, Monica Liesch, Urs Keller
Methodology: Judith V Graser, Hubertus JA van Hedel, Laura Prospero, Monica Liesch
Project administration: Judith V Graser, Laura Prospero, Monica Liesch
Resources: Hubertus JA van Hedel
Validation: Judith V Graser, Urs Keller, Hubertus J. A. van Hedel
Visualization: Judith V Graser
Writing – original draft: Judith V Graser
Writing – review & editing: Judith V Graser, Hubertus JA van Hedel, Urs Keller

ADDITIONAL FILES

Additional file 3.1.

A



(Continued)

Additional file 3.1. Continued

B

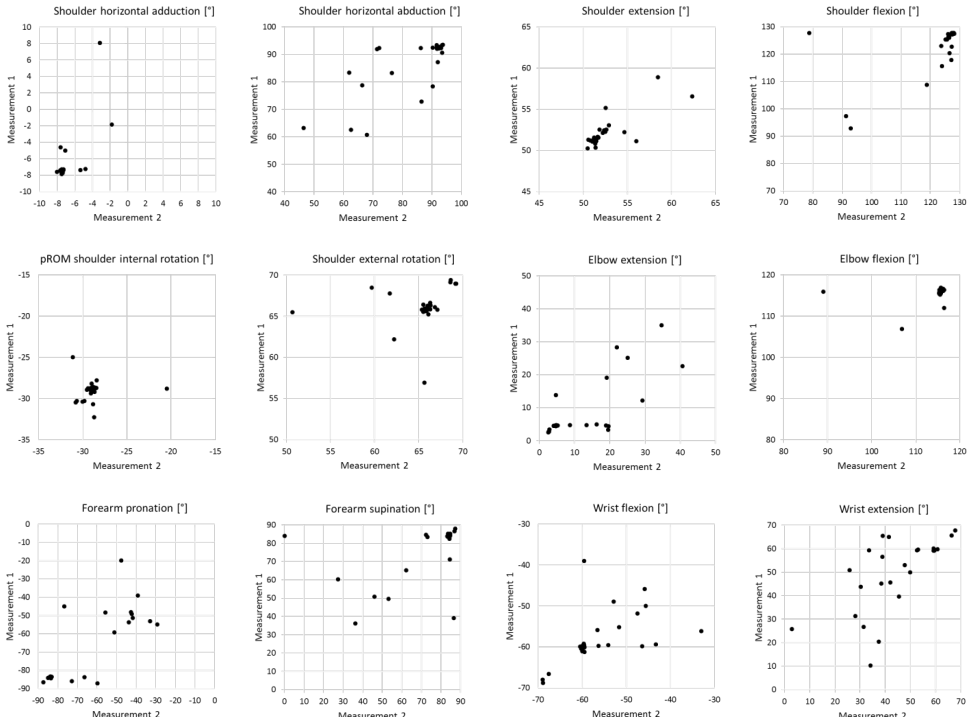
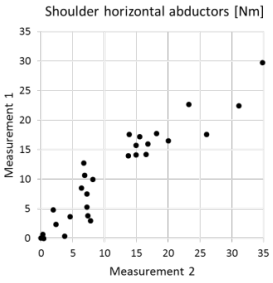
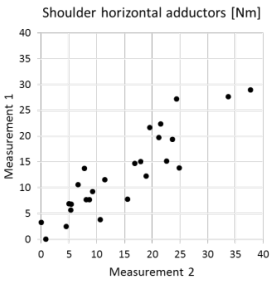
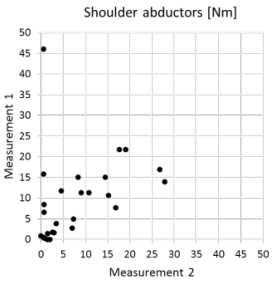


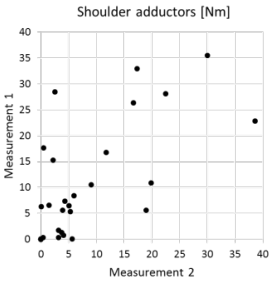
Fig. A3.1. Distribution of the data of each parameter obtained from the **active and passive Range of Motion assessments**. Displayed are the data of each parameter of A) active Range of Motion (aROM) assessment and B) passive Range of Motion assessment (pROM). Each movement axis contains two movement directions. Movement directions are displayed as negative or positive numbers, respectively: shoulder horizontal adduction and abduction, shoulder extension and flexion, shoulder internal rotation and external rotation, elbow extension and flexion, forearm supination and pronation, wrist flexion and extension. The respective end-position of the joint angle is displayed in degrees (°). The X-axis represents the second measurement, the Y-axis represents the first measurement.

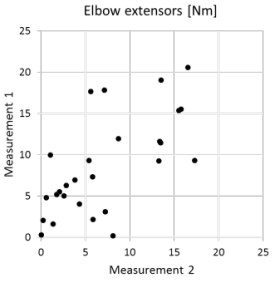
Additional file 3.2.

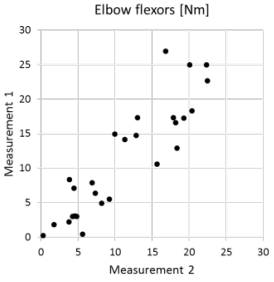


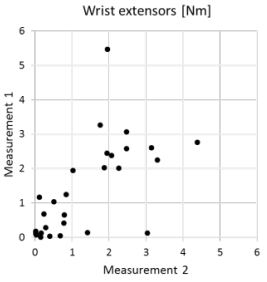


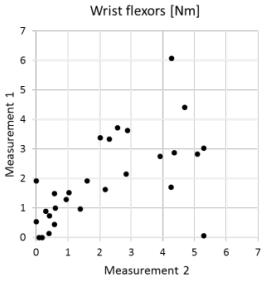












(Continued)

Additional file 3.2. Continued

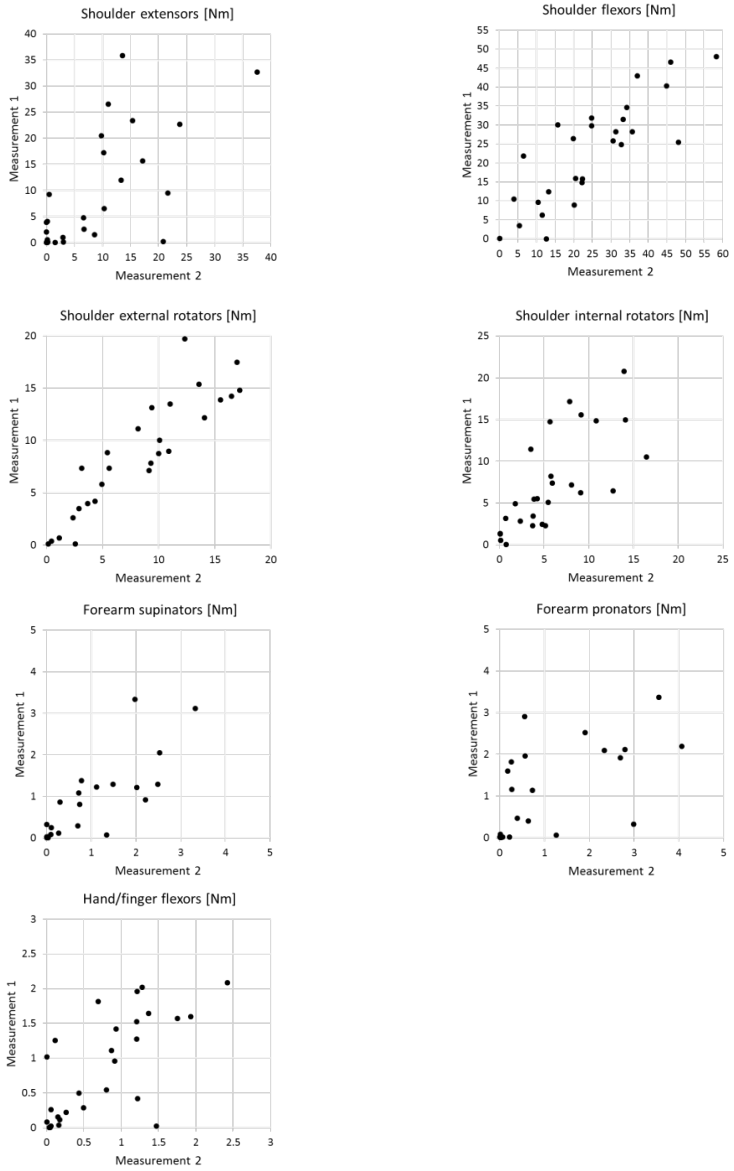
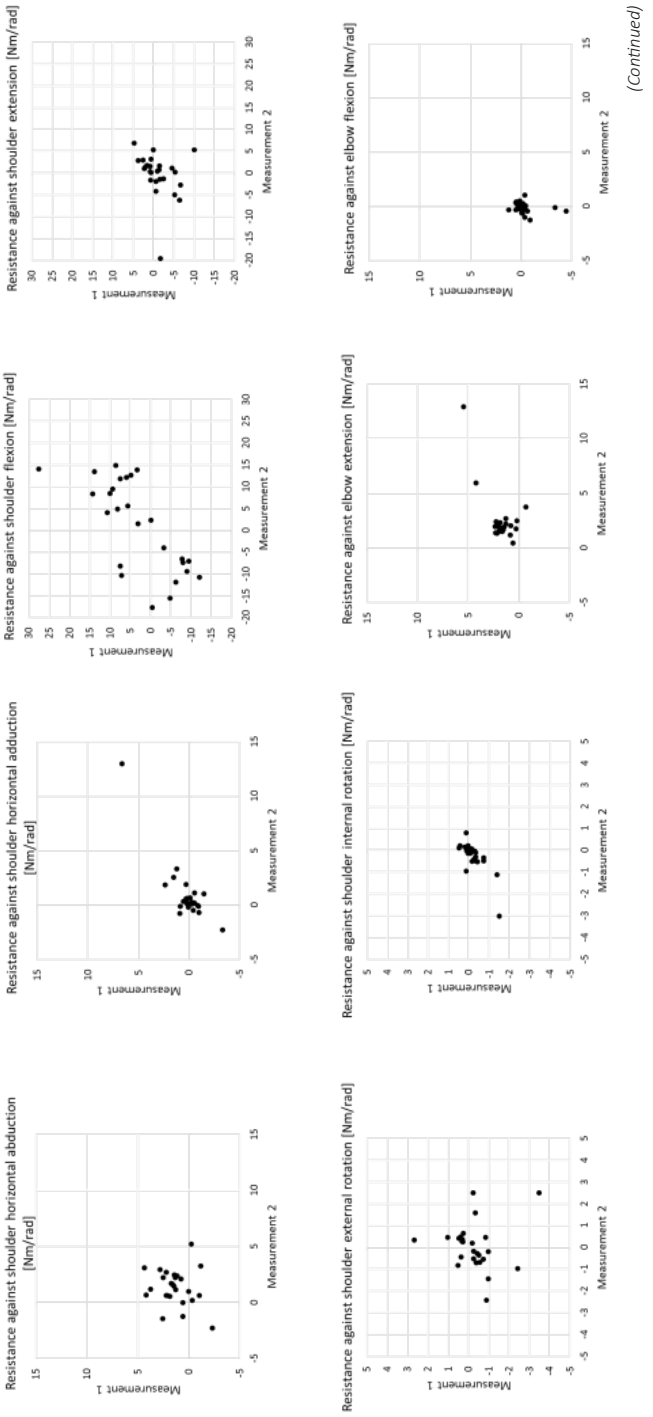


Fig. A3.2. Distribution of the data of each parameter obtained from the **Strength assessment**. Displayed are the absolute data of each parameter the Strength assessment in Newton meter (Nm): Strength of the shoulder horizontal abductors and adductors, shoulder extensors and flexors, shoulder abductors and adductors, shoulder external rotators and internal rotators, elbow extensors and flexors, forearm supinators and pronators, wrist extensors and flexors and hand/finger flexors. The X-axis represents the second measurement, the Y-axis represents the first measurement.

Additional file 3.3



(Continued)

Additional file 3.3 Continued

Fig. A3.3. Distribution of the data of each parameter obtained from the **Resistance to Passive Movement assessment**. Displayed are the data of each parameter the Resistance to Passive Movement (RPM) assessment in Newton meter per degrees (Nm/rad). The plots show the respective differences between the resistance against passive movements with a speed of 60 seconds per degree and the resistance against passive movements with a speed of 10 seconds per degree of the following movement directions: resistance against horizontal shoulder abduction and adduction, shoulder flexion and extension, shoulder external and internal rotation, elbow extension and flexion, forearm supination and pronation, wrist extension and flexion. The X-axis represents the second measurement, the Y-axis represents the first measurement.

Additional file 3.4

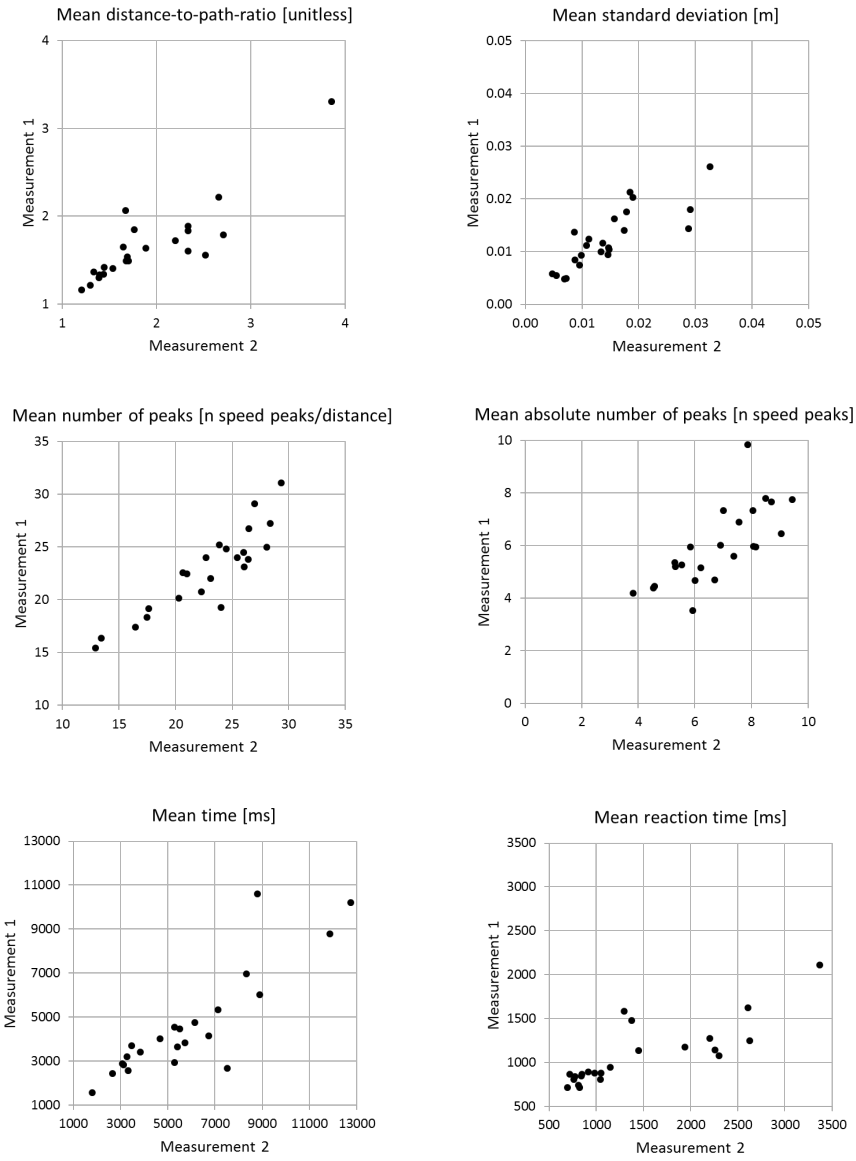


Fig. A3.4. Distribution of the data of each parameter obtained from the **Quality of Movement assessment**. Displayed are the data of each parameter the Quality of Movement (QoM) assessment (means of three trials): The mean distance-to-path-ratio (unitless), the mean standard deviation on the target in meters (m), the mean number of speed peaks per covered distance (n speed peaks/distance), the mean absolute number of speed peaks (n speed peaks), the mean time needed to reach a target, in milliseconds (ms) and the mean reaction time in milliseconds (ms). The X-axis represents the second measurement, the Y-axis represents the first measurement.

Additional file 3.5

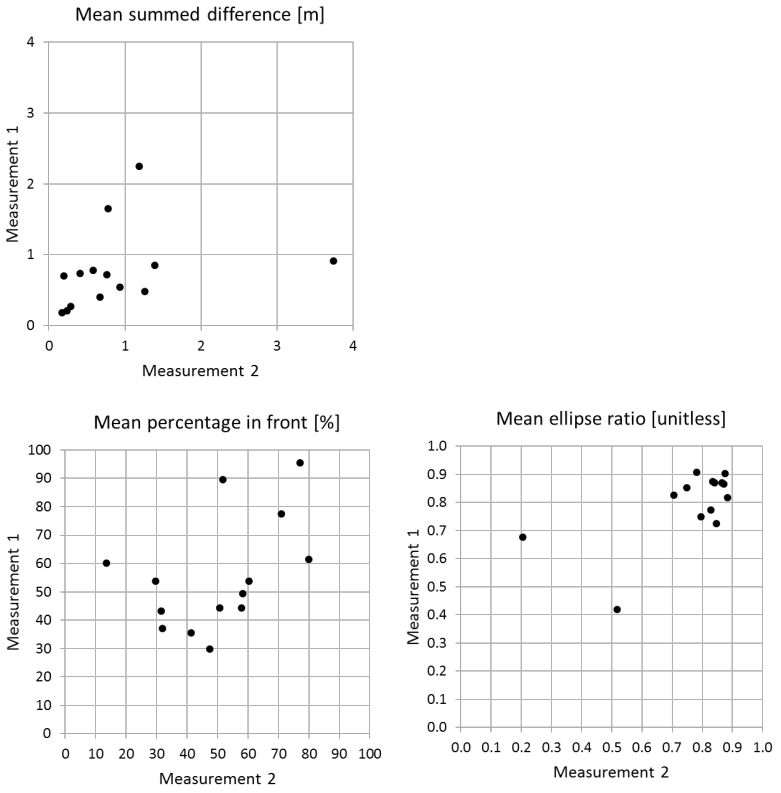
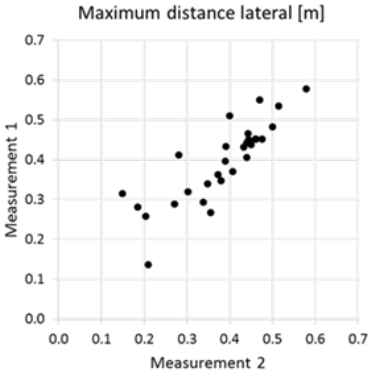
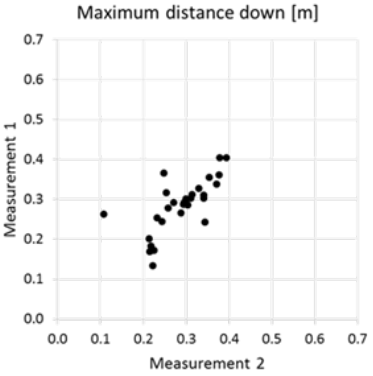
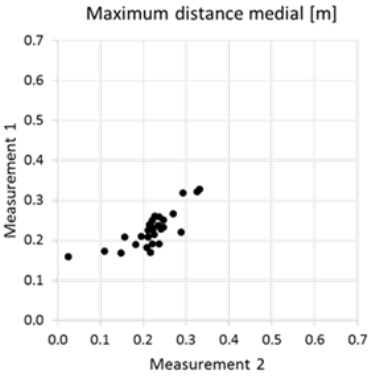


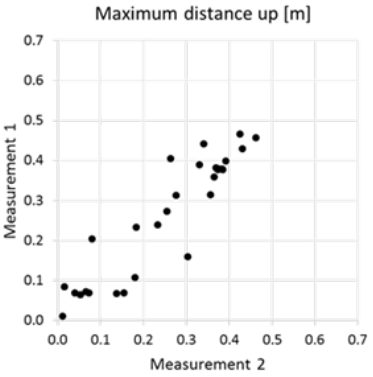
Fig. A3.5. Distribution of the data of each parameter obtained from the **Circle assessment**. Displayed are the data of each parameter the Circle assessment (means of three trials): Mean summed difference between the current position and the reference circle in meters (m), mean percentage of the time the participant was in front of the reference circle in percentage (%), and the mean ellipse ratio between the minimal and the maximal radius of the least-squares fitted ellipse (unitless). The X-axis represents the second measurement, the Y-axis represents the first measurement.

Additional file 3.6



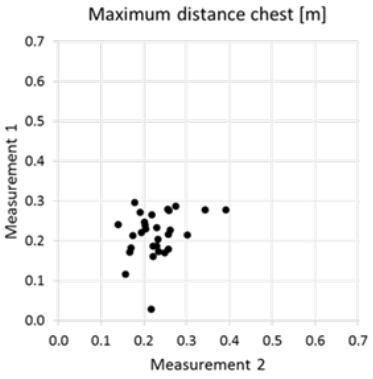


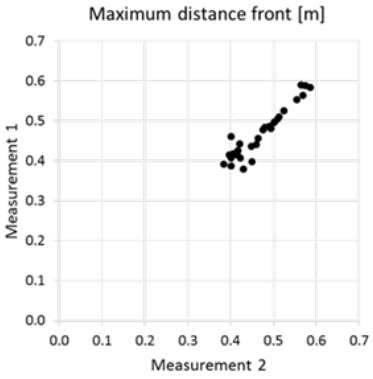




(Continued)

Additional file 3.6 Continued





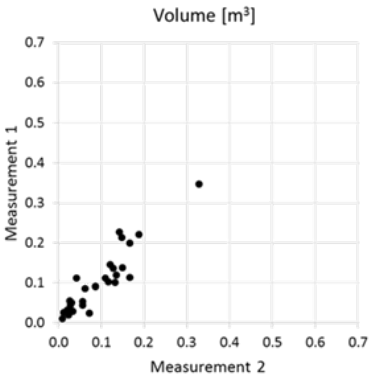


Fig A3.6. Distribution of the data of each parameter obtained from the **Workspace assessment**. Displayed are the absolute data of each parameter the Workspace assessment: maximum distance in medial direction, lateral direction, down, up, to the front and to the chest in meters (m) and the maximum reachable volume in cubic meters (m3). The X-axis represents the second measurement, the Y-axis represents the first measurement.

REFERENCES

1. Shikako-Thomas K, Majnemer A, Law M, Lach L. Determinants of participation in leisure activities in children and youth with cerebral palsy: Systematic review. Phys Occup Ther Pediatr. 2008;28:155–69.
2. Makki D, Duodu J, Nixon M. Prevalence and pattern of upper limb involvement in cerebral palsy. J Child Orthop. 2014;8:215–9.
3. Bayón C, Raya R, Lerma Lara S, Ramírez Ó, Serrano J I, Rocon E. Robotic therapies for children with cerebral palsy: A systematic review. Transl Biomed. 2016;7:1–10.
4. Bryanton C, Bossé J, Brien M, McLean J, McCormick A, Sveistrup H. Feasibility, motivation, and selective motor control: Virtual reality compared to conventional home exercise in children with cerebral palsy. Cyberpsychol Behav. 2006;9:123–8.
5. Jannink MJA, Van der Wilden GJ, Navis DW, Visser G, Gussinklo J, Ijzerman M. A low-cost video game applied for training of upper extremity function in children with cerebral palsy: A pilot study. Cyberpsychology Behav. 2008;11:27–33.
6. Holden MK. Virtual environments for motor rehabilitation: Review. Cyberpsychology Behav. 2005;8:187–211.
7. Keller JW, van Hedel HJ. Weight-supported training of the upper extremity in children with cerebral palsy: A motor learning study. J Neuroeng Rehabil. 2017;14.
8. Falzarano V, Marini F, Morasso P, Zenzeri J. Devices and protocols for upper limb robot-assisted rehabilitation of children with neuromotor disorders. Appl Sci. 2019;9.
9. Meyer-Heim A, van Hedel HJA. Robot-assisted and computer-enhanced therapies for children with cerebral palsy: Current state and clinical implementation. Semin Pediatr Neurol. 2013;20:139–45.
10. Cimolin V, Germiniasi C, Galli M, Condoluci C, Beretta E, Piccinini L. Robot-assisted upper limb training for hemiplegic children with cerebral palsy. J Dev Phys Disabil. 2019;31:89–101.
11. Leem MJ, Kim GS, Kim KH, Yi TI, Moon HI. Predictors of functional and motor outcomes following upper limb robot-assisted therapy after stroke. Int J Rehabil Res. 2019;42:223–8.
12. Shin J, Kim M-Y, Lee J-Y, Jeon Y-J, Kim S, Lee S, et al. Effects of virtual reality-based rehabilitation on distal upper extremity function and health-related quality of life: A single- blinded, randomized controlled trial. J Neuroeng Rehabil. 2016;13.
13. Keller U, Van Hedel HJA, Klamroth-Marganska V, Riener R. ChARMin: The first

actuated exoskeleton robot for pediatric arm rehabilitation. IEEE/ASME Trans Mechatronics. 2016;21:2201–13.

14. Schwarz A, Kanzler CM, Lambercy O, Luft AR, Veerbeek JM. Systematic review of upper limb movements after stroke. Stroke. 2019;50:718–27.
15. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Ohrvall A-M, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: Scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006;48:549–54.
16. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther. 1987;67:206–7.
17. Trompetto C, Marinelli L, Mori L, Pelosin E, Currà A, Molfetta L, et al. Pathophysiology of spasticity: Implications for neurorehabilitation. Biomed Res Int. 2014;2014.
18. Weir JP. Quantifying Test-Retest Reliability Using the Intraclass Correlation Coefficient and the SEM. J Strength Cond Res. 2005;19:231–40.
19. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15:155–63.
20. Blanca MJ, Alarcón R, Arnau J, Bono R, Bendayan R. Non-normal data: Is ANOVA still a valid option ? Psicothema. 2017;29:552–7.
21. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine. A Practical Guide. Cambridge: Cambridge University Press; 2011.
22. Liljequist D, Elfving B, Roaldsen KS. Intraclass correlation – A discussion and demonstration of basic features. PLoS One. 2019;14.
23. Colaris J, van der Linden M, Selles R, Coene N, Allema JH, Verhaar J. Pronation and supination after forearm fractures in children: Reliability of visual estimation and conventional goniometry measurement. Injury. 2010;41:643–6.
24. Klingels K, De Cock P, Molenaers G, Desloovere K, Huenaerts C, Jaspers E, et al. Upper limb motor and sensory impairments in children with hemiplegic cerebral palsy. Can they be measured reliably? Disabil Rehabil. 2010;32:409–16.
25. Zhang M, Zhang S, McDaid A, Davies C, Xie SQ. Automated objective robot-assisted assessment of wrist passive ranges of motion. J Biomech. 2018;73:223–6.
26. Kendall FP, Kendall Mc Creary E, Geise Provance P. Muscles— testing and function. 4th editio. Philadelphia: Williams and Wilkins; 1993.
27. Lance JW. Pathophysiology of spasticity and clinical experience with Baclofen.

In: Lance JW, Feldman RG, Young RR, Koella WP, editors. Spasticity: Disordered motor control. Chicago: Year Book; 1980. p. 185–204.

28. Fleuren J, Voerman G, Erren-Wolters C, Snoek G, Rietman J, Hermens H, et al. Stop using the Ashworth Scale for the assessments of spasticity. J Neurol Neurosurg psychiatry. 2010;81:46–52.

29. Knudson D. The biomechanics of stretching. J Exerc Sci Physiother. 2006;2:3–12.

30. Skjaerven LH, Kristoffersen K, Gard G. An eye for movement quality: A phenomenological study of movement quality reflecting a group of physiotherapists' understanding of the phenomenon. Physiother Theory Pract. 2008;24:13–27.

31. Keller U, Schölch S, Albisser U, Rudhe C, Curt A, Riener R, et al. Robot-assisted arm assessments in spinal cord injured patients: A consideration of concept study. PLoS One. 2015;10.

32. Finley MA, Dipietro L, Ohlhoff J, Whitall J, Krebs HI, Bever CT. The effect of repeated measurements using an upper extremity robot on healthy adults. J Appl Biochechanics. 2009;25:103–10.

33. Rudhe C, Albisser U, Starkey ML, Curt A, Bolliger M. Reliability of movement workspace measurements in a passive arm orthosis used in spinal cord injury rehabilitation. J Neuroeng Rehabil. 2012;9.

34. van Hedel HJA, Lieber J, Ricklin S, Meyer-Heim A. Die praktische Anwendung von Exergames und virtueller Realität in der pädiatrischen Rehabilitation. neuroreha. 2017;09:35–40.

35. Randall M, Johnson L, Reddihough D. The Melbourne Assessment 2. https://www.rch.org.au/melbourneassessment/. Accessed 29 Mar 2019.

36. Krumlinde-Sundholm L, Eliasson A-C. Development of the Assisting Hand Assessment: A Rasch-built measure intended for children with unilateral upper limb impairments. Scand J Occup Ther. 2009;10:16–26.

CHAPTER 4

CONTEXTUAL INTERFERENCE WITH CHARMIN EXERGAMES IN CHILDREN WITH BRAIN LESIONS AND IMPAIRED UPPER LIMB FUNCTION - A FEASIBILITY EXPERIMENT

Unpublished chapter

ABSTRACT

Introduction: Evidence about the contextual interference is missing in children with brain lesions. We therefore aimed to conduct a feasibility study about the contextual interference effect in children with congenital or acquired brain lesions and impaired upper limb function when practicing two exergames with the ChARMin exoskeleton. Our objectives were to explore the acceptability, the implementation and the practicality of the study procedures.

Methods: We recruited children with congenital or acquired brain lesions undergoing an inpatient rehabilitation stay. We randomly allocated them to either a blocked or a random order group. Depending on their allocation, they performed one session of two exergames with the ChARMin exoskeleton for upper limb training either in blocked or in random order during the second to last therapy session of the day. The next morning, the same session was repeated. Assessments (Box and Block Test, Quality of Movement Assessment and Circle following, provided by ChARMin) were performed before, immediately after, and one hour after the first practice session, before and immediately after the second practice session. Data were collected by systematically taking notes throughout the study appointments, which were then categorised into 'acceptability of the participants', 'implementation of the intervention in the clinical context', and 'practicality of the study procedure'.

Results: The main issue was the technical problems of the ChARMin device. The compliance of the participants was closely related to the number and severity of the technical problems. Furthermore, the implementation of the study and its single components needs to be revised and adapted for a future study.

Conclusion: This straight-forward and explorative feasibility approach revealed that extensive adaptations are required. A more detailed study protocol with clearly defined primary outcome and feasibility criteria is needed to enable a clear decision about the feasibility of a future trial investigating the contextual interference effect of upper limb motor tasks in children with congenital or acquired brain lesions using the ChARMin exoskeleton robot.

INTRODUCTION

Contextual interference predicts that practicing in random order results in better performance during retention of the learned skill and when transferring it to another skill, compared to blocked order which leads to better acquisition of the skill [1]. Yet, the contextual interference effect has mainly been shown in healthy adults and while knowledge about it has been considered important also in paediatric neurorehabilitation, our systematic review has shown a lack of evidence about the hypothesized effect in this field [2]. According to our knowledge, there is no motor learning study comparing random versus blocked practice order in children with brain impairments using robotic devices. New technologies come with the opportunity to measure and train motor functions in a standardised way and provide playful and highly repetitive training through exergames with both functions implemented in the same technology. Hence, using a robotic device for the investigation of the contextual interference effect in children with neuromotor impairments seemed promising. We chose ChARMin, an exoskeleton designed to both practice and measure upper limb functions [3] to perform our first exploratory experiments. It is suggested that ChARMin can be applied in treatment sessions with children aged five years and older due to the two different distal modules and especially in children with severe impairments [3]. Because of these features and because upper limb functions can be trained playfully with exergames, we decided that ChARMin would be a suitable device for this experiment.

A real challenge would be to design a study that includes a valid comparison between interventions with a high versus low level of contextual interference. A prerequisite for the implementation for such a design is that it should be manageable for both, participants and researchers. With so many open questions, and no experience with using ChARMin in an interventional study, the conduct of a feasibility experiment with a limited number of participants seemed a necessary first step. The main objective was to gain knowledge about the criteria that are necessary to perform a randomised controlled experiment on this topic within the context of our rehab center including the patients.

Specific aims of this small feasibility study were to explore 1) the participants' acceptability of the study procedures; 2) the implementation of the intervention in the clinical context; and 3) the practicality of the study procedure.

METHODS

Study design

For this exploratory small-scaled experiment, we planned a randomised controlled parallel group design, evaluating practice of two exergames with the ChARMin robot in either blocked or random order.

Recruitment

Inpatients undergoing neurorehabilitation at the Swiss Children's Rehab, University Children's Hospital Zurich, Affoltern am Albis, Switzerland, were informed about the study. When agreeing verbally, written informed consent was obtained from the participants and the parents or authorised representatives This study has been approved by the Ethics Committee of the Canton Zurich (BASEC-Nr. PB_2016-02450) and the Swiss Agency for Therapeutic Products (Swissmedic reference number: 2015-MD-0009).

Participants

Included were children meeting the following inclusion criteria: age 5 to 18 years old, diagnosis of congenital or acquired brain lesion with impaired upper limb function (unilateral or bilateral), Manual Ability Classification System (MACS) level I-IV (i.e. at least the handling of a limited selection of easily manageable objects in adapted situations should be possible), the ability to sit upright for approximately 60 minutes without lateral trunk support, the ability to follow instructions, the ability to report discomfort and/or pain and a visual ability sufficient to follow the events on the computer screen. Exclusion criteria were skin lesions on the tested upper limb, upper-limb surgery or Botox injections during the past six months and further exclusion criteria determined by the ChARMin manual of operations [4].

Sample size justification

As we collected qualitative data regarding the experimental procedure, we started with the measurements and evaluated the experiences of relevant stakeholders (i.e. participants and therapists). We would recruit and measure until we would have gained enough information (data saturation) about the objectives. Hence, no sample size was determined.

Randomisation

The principle investigator allocated participants to the blocked or random order group using a predefined electronically generated randomisation key. According to this key, we prepared a list relating the participant's ID number with the blocked or random order practice group.

Study procedure

The study procedure is displayed in Figure 4.1. After recruitment and random allocation of the participants, baseline measurements were performed. All the following sessions took place on two consecutive days. As the study sessions had to be fitted within the planning structure of the rehabilitation programme of our centre, which has defined therapy slots throughout the day, we planned the sessions as follows: Practice session 1 was performed during the second to last therapy slot of the first day and was preceded and followed by an assessment block. The last

slot of the day was planned as a break with no therapies, which was followed again by an assessment block. The idea was to prevent a potential influence on the last assessment block by any therapies. We also asked the participants not to play any computer games during this break. The second practice session was performed during the first therapy slot the next morning and was again framed by two assessment blocks.

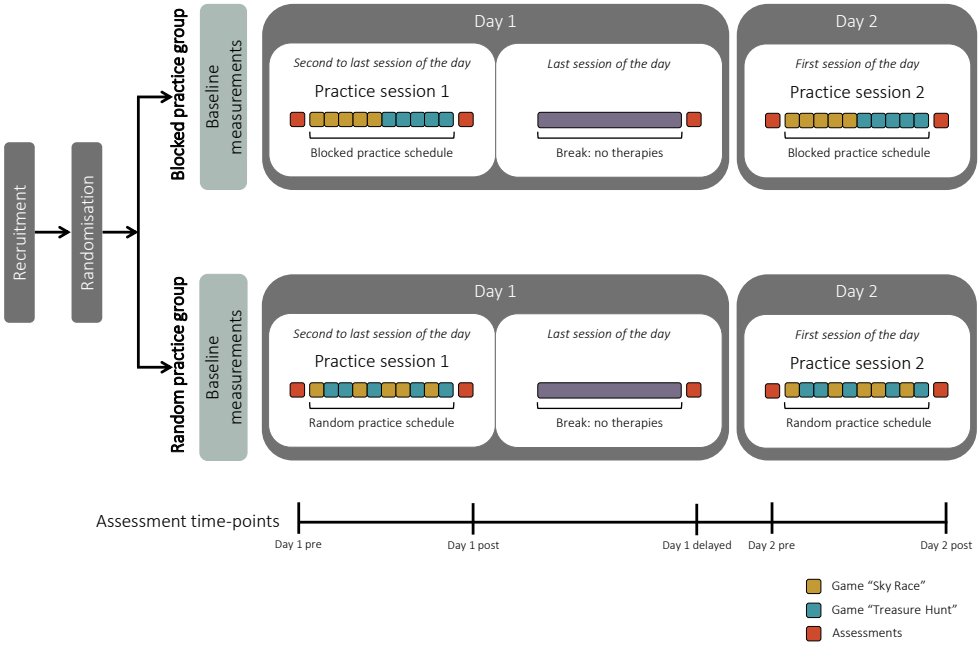


Fig. 4.1. Study procedure. After recruitment and randomisation, baseline measurements such as the Manual Ability Classification System (MACS) or Test of Nonverbal Intelligence, version 4 (TONI-4) were conducted. Participants performed two practice sessions either in blocked or random order on two subsequent days. Assessments were performed before and after each practice session and one hour after the first practice session.

Device, exergames and practice

ChARMin (Figure 4.2), an exoskeleton for upper limb training, has been developed in a collaboration between the Sensory-Motor Systems Lab (Federal Institute of Technology, Zurich, Switzerland) and the Swiss Children's Rehab of the University Children's Hospital Zurich.

ChARMin comes with two distal modules of different sizes (small for younger children and large for older children). It provides three different assistance modes (fully-guided, assisted as needed and no assistance). The level of support against gravity can be adjusted according to the child's ability, however, in this experiment, we applied neither assistance nor support against gravity. Movements involving

the whole upper limb (shoulder, elbow, and wrist) or only single joints are used to play motivating exergames to practice several upper limb functions and perform assessments to measure these functions [5].

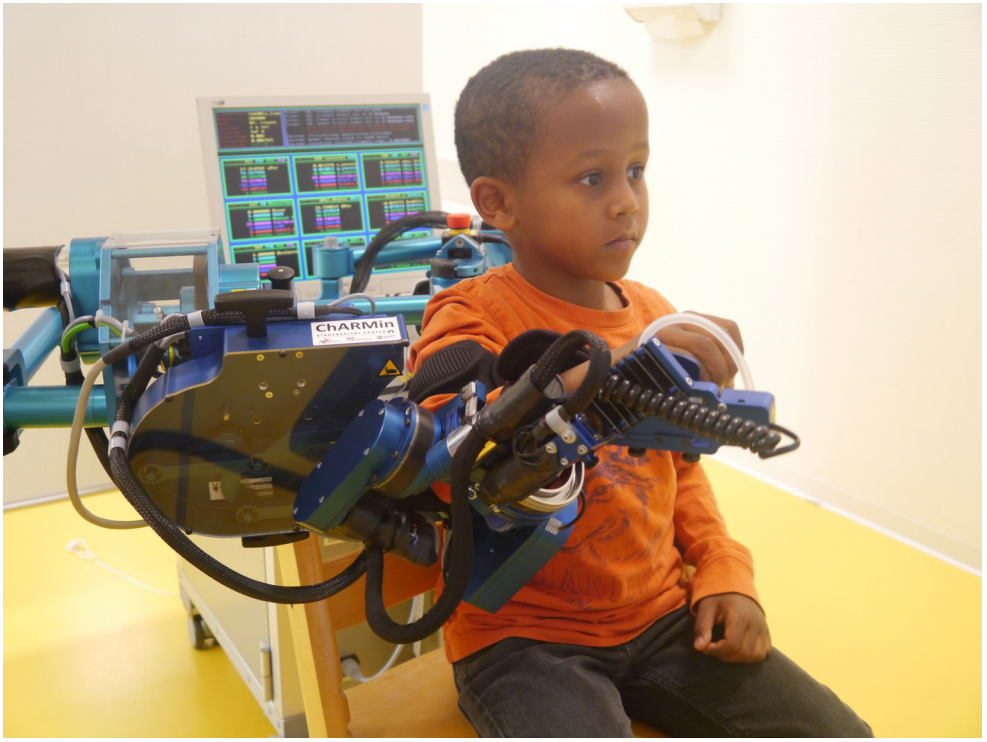


Fig. 4.2. The ChARMin exoskeleton. A child practicing upper limb functions by playing exergames with the ChARMin.

The two exergames used for the practice sessions in this feasibility study are Sky Race and Treasure Hunt (Figure 4.3). Sky Race is an exergame in which the participant is steering either a plane or a bird with his arm through a fantasy world of a city whose houses are flying through the air. The aim is to avoid crashing into the obstacles (houses, red balloons) and to collect as many 'good' objects (yellow balloons and rings, red hearts) as possible. In Treasure Hunt, the participant is a diver walking through an underwater world and is required to aim at different objects by moving the arm. The aim is to shoot the 'bad' pirate fish and to collect coins hidden in treasure boxes. The game scores of both exergames are based on the number of collected objects. Points are deducted for crashing in houses and red balloons (Sky Race) and for shooting the 'good' fish (Treasure Hunt).

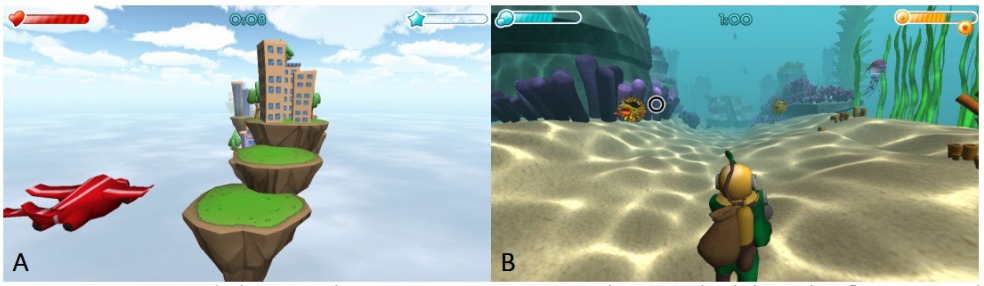


Fig. 4.3. Exergames. A) Sky Race: the participants steer an airplane or a bird through a flying city and need to avoid obstacles and catch several objects. B) Treasure Hunt: a diver walks on the ground of the ocean, the participant has to catch treasure chests and pirate fish.

We chose these exergames because different kinds of movements are required to play them. While for Sky Race, the majority of the movements are more or less constant (steering a flying object) in Treasure Hunt, as the diver does not need to be steered, the movements are rather prone to reach targets, resulting in single goal directed movements. We considered these differences as suitable in terms of using them as variations for the different practice orders to generate different levels of contextual interference when practicing these exergames either in blocked or in random order. The exergames were performed with the more affected side. If both sides were similarly affected, the dominant side was chosen.

The numbers of trials (one-minute duration each) were established bit by bit during the practice sessions in both practice groups. Depending on the physical ability and the child's motivation, we determined the possible number of trials individually for each child. We aimed to apply the same number of trials in both sessions and exergames.

The blocked order group practiced the full block of trials of the first exergame before switching to the other exergame. The order of the exergames (i.e. whether participants first practiced a block of Treasure Hunt or Sky Race) was randomised. Participants in the random order group performed the trials of the exergames in a random order with the exception that the same exergame could not be played more than twice in a row to keep the contrast between the two practice orders as large as possible.

Measures

Baseline measures

To characterise the participants, we collected clinical data from patient records for the following parameters: age, gender, body height and weight, diagnosis, and the Gross Motor Function Classification System level. The level of the Manual Ability Classification System (MACS) [6] was obtained for each participant either from the patient record or the responsible occupational therapist. The MACS describes how

children use their hands in activities of daily life. We included children with MACS level I (handling objects easily and successfully) to level IV (handles a limited selection of easily managed objects in adapted situations).

The Melbourne Assessment 2 (MA2) [7] was performed by an experienced and trained occupational therapist. The MA2 quantifies the quality of unilateral upper limb motor function in the four subscales range of motion, fluency, accuracy, and dexterity [8]. The MA2 is reliable (Intraclass Correlation Coefficients (ICCs) in all four subscales 0.92-0.98), valid (correlating moderately to strongly (Pearson correlation coefficients of 0.40-0.79) with the subtest 3 of the Bruininks-Oseretsky Test of Motor Proficiency, 2nd Edition, the Box and Blocks Test (BBT), and the Pediatric Motor Activity Loge Revised), and responsive (standardised response means 0.70 – 2.00 for all the four subscales) when applied in children with cerebral palsy [8].

The Test Of Nonverbal Intelligence (Version 4, TONI-4) [9] was supervised and interpreted by a neuropsychologist. It evaluates abstract reasoning and problem solving (Index Score <70: very poor, 70-79: poor, 80-89: below average, 90-110: average, 111-120: above average, 121-130: superior, and >130: very superior) [9].

Feasibility measures

Exploring the feasibility, the focus lay on 1) the acceptability of the study procedures as demonstrated by the compliance and motivation of the participants; 2) the implementation of the intervention in the clinical context (e.g. planning and organising the research appointments within the rehabilitation schedule); the experience of the therapist when conducting the interventions; and 3) the practicality, which included aspects such as understanding and following the instructions, having the physical ability to perform the exergames and assessments, and the use of the ChARMin. We took notes during the appointments.

Motor learning measures

Acquisition: To measure the performance during the acquisition (while playing the exergames), we decided to use the game scores provided by the exergames themselves. They reflect the efficacy of the goal-directed movement in terms of reaching and avoiding certain points (targets) on the screen. Behind this efficacy stands a motor response time to a stimulus (appearing of target) and movement speed as targets can be missed when not reacting fast enough.

Retention: For evaluating whether the participants were able to retain the practiced skills, we also chose the game scores obtained while playing the exergames at the beginning of session 2 compared to the end of session 1.

Transfer: To measure a motor learning transfer, we chose the Quality of Movement (QoM), and Circle following (Circle) assessments provided by ChARMin, and the BBT [10]. During QoM, the participants were instructed to reach eight targets appearing

radially around the centre point of the screen (Figure 4.4). A new target appeared when the centre point was reached again. QoM records parameters reflecting movement quality of goal-directed movements (i.e. the mean distance-to-path ratio, the mean standard deviation on the target, the mean number of velocity peaks, (absolute and normalised to the actual distance covered), the mean time needed for the movement, and the mean reaction time) [5]. Our evaluations in children with congenital or acquired brain lesions have shown that the parameters obtained with QoM are moderately to very highly reliable (test-retest reliability) with ICCs ranging from 0.69 to 0.94 [5]. The Smallest Real Differences as a percentage of the grand means (SRD%) showed rather large measurement errors from 17.71% to 81.26% [5]. During Circle, the participants were asked to follow the predetermined circle-movement by staying on a green ball moving on this given path (Figure 4.4). During the movement, the tracking capabilities of the hand are recorded (parameters: the mean summed difference between the current position and the reference circle, the mean percentage moving in front of the green ball, the ratio between the minimal and the maximal radius of the least-squares fitted ellipse) [5]. The parameters of Circle showed low to high test-retest reliability (ICCs ranging from 0.42 to 0.73) and large measurement errors (SRD% ranging from 37.3% to 218.1%) in children with either congenital or acquired brain lesions [5]. As we did not have a feasible alternative to measure quality of movement and since QoM as well as Circle would provide parameters reflecting what we supposed would be practiced during the two exergames (namely, reaching targets, reaction time, movement speed etc.), we decided to apply these assessments despite their rather low reliability. Having an idea about the measurement errors would enable us to interpret the results accordingly. Both, QoM and Circle assessments were repeated three times each.

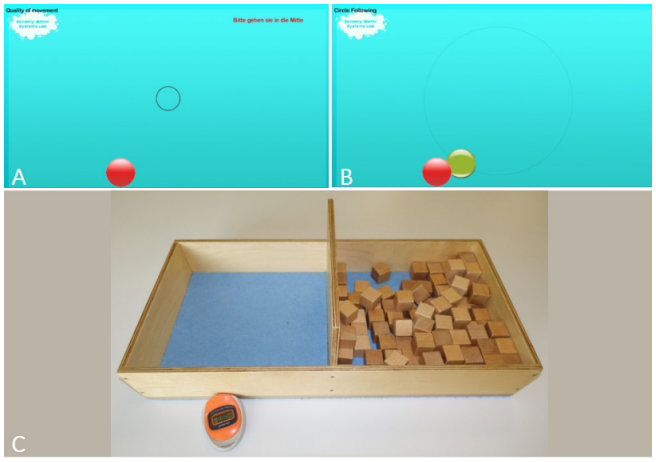


Fig. 4.4. Assessments. A) Circle, measures tracking capabilities of the hand, B) Quality of Movement measures movement quality of goal-directed movements, C) Box and Blocks Test assesses manual dexterity.

Additionally, we chose the BBT as non-robotic assessment, which can be easily applied also because it is not time-consuming. The BBT assesses the gross manual dexterity [10]. The participant has to pick wooden blocks out of a compartment of a box and transport them into the other compartment. The aim is to pick and transport as many blocks within one minute as possible (Figure 4.4). The BBT has been described as valid and reliable in typically developing children, at the age between three and ten years (concurrent validity: correlations with the Movement Assessment Battery for Children–2 from -0.25 to -0.72, test-retest reliability ICC = 0.85, interrater reliability ICC = 0.99) [11]. Originally, the BBT is used for testing the dominant and the non-dominant hand and participants are instructed to perform the test once with each hand [12]. As we were only interested in the hand with which the children practiced the exergames, the BBT was performed twice and we used the average value for the analyses.

Data analysis

Feasibility data

The notes taken during the different phases of the study were coded with different colours to be allocated to one of the following topics: (1) acceptability: participants; (2) implementation: planning procedure, therapist; (3) practicality: instructions, physical ability, use of ChARMin. These notes were typed down. Notes that could not be allocated to one of the topics were added under the category 'comments'. The study team consisting of the person in charge of recruitment, planning, conducting appointments, and gathering data, the leader of the research group and the engineer who developed ChARMin conducted group discussions about the findings. Based on the notes, experiences that were reported during these discussions, and information based on literature, we would determine further proceedings and adaptation of the study design.

Motor learning data

It was not our objective to get information about motor learning with this exploratory experiment as the focus was on the feasibility of the procedure. The following considerations about the analysis of motor learning served the understanding of how acquisition, retention and transfer could be looked at in the future. We performed individual data exploration.

We defined acquisition as the performance throughout the two practice sessions (comparison of the mean of game scores of first and last three trials per session), retention as the difference of the mean game score at the end of session 1 and the beginning of session 2. Immediate transfer was defined as the performance during the BBT, QoM, and Circle immediately after sessions 1 and 2. Delayed transfer was defined as the performance in these three assessments one hour after finishing session 1. All comparisons would be assessed for each game individually.

RESULTS

Three participants were recruited and randomly allocated to the blocked or the random group. Participant characteristics obtained at baseline are presented in Table 4.1.

Feasibility

1. **Acceptability:** Two participants had problems with compliance. This became obvious as rules of the exergames were disregarded on purpose and a lot of verbal motivation was needed from the therapist. We aimed to apply an individual total number of repetitions per session for each participant, equally distributed to both exergames. The second session was supposed the have the identical sequence. Yet, we could not adhere to this procedure strictly because some sessions had to be aborted ahead of schedule. Table 4.2 shows the number of trials per exergame and session. In one case, the reason for the abortions was the time limit of the session that was reached (end of the therapy slot). In the other cases, the reasons were lack of compliance, motivation and/or fatigue of the participants.

2. **Implementation:** Experience about the planning procedure was positive as every appointment could be planned despite the high number of fixed appointments that each child already had in their regular rehabilitation schedule. Yet, when discussing the results with the study team, concerns about whether these two different exergames would generate a contextual interference effect were raised. This issue is further discussed in the discussion section of this chapter. Therapists experienced some issues related to the rather unfamiliar implementation of the robot during the intervention. The procedure (assessments/exergames) needed to be interrupted because the robot settings needed to be adapted (e.g. the safety pedal was not activated, a wrong speed was chosen for an exergame, or the wrong profile name was entered).

3. **Practicality:** The participants understood the instructions well and followed them. Only one participant needed additional instructions on two occasions. Concerning the physical ability of the participants, we noted that during the practice phase of day 1, one participant mentioned that the procedure was exhausting and he needed a short break. This participant also needed additional help from the therapist during the pre-assessments on day 1. On day 2, it was mentioned that weight support was needed to enable the participant to perform the sessions. Otherwise, reaching the upper areas on the screen would not have been possible. Another participant helped himself with the free arm. It is unclear whether this was due to exhaustion or not.

There was one note about a longer break (35 minutes) during a delayed assessment on day 1 (after the BBT). Unfortunately, the reasons for this break were not noted. It

could have been necessary to take a break due to fatigue of the participant or lack of compliance, or also due to technical problems when starting the ChARMin.

The highest number of notes were taken on the use of ChARMin. We noted 21 error messages (e.g. when the end effector speed was too high or the absolute movement limit was reached in one of the axes). Most of these errors required new starts of the games or assessments. These errors were the main reason for incomplete datasets. Movement oscillations (when the movement of the patient's arm is amplified by the active joint friction compensation of the robot, there is an overshoot of this movement, leading to these oscillations) of the robotic arm (26 times noted) were also a major issue. While the assessment or exergame could be continued, data quality would be affected.

Table 4.1. Participant characteristics

ID	Age [yrs]	Gender	Body height [cm]	Body weight [kg]	Diagnosis	GMFCS level	MACS level	MA2 total % score [%]	TONI-4 index score
1	11.2		138.5	37.5	CP bilateral, spastic, right side more affected	II	II	85.4	85
2	7.8		121.0	22.7	CP bilateral, spastic, lower limbs and right side more affected	II	I	95.5	109
3	11.6		142.3	29.4	CP unilateral, spastic, right side more affected	II	II	33.7	84

Abbreviations: ID = identification, CP = cerebral palsy, GMFCS = Gross Motor Function Classification System, MACS = Manual Ability Classification System, MA2 = Melbourne Assessment 2, TONI-4 = Test Of Nonverbal Intelligence.

Table 4.2. Practice characteristics

ID	Practice order	Practice session 1	Practice session 2
		n trials Sky Race/Treasure Hunt	n trials Sky Race/Treasure Hunt
1	Random	17/18	8/8
2	Blocked	10/10	15/15
3	Blocked	11/10	11/10

Abbreviations: ID = identification, n = number

Motor learning

Acquisition and retention

For the acquisition, we evaluated the game score recorded for each trial. By taking the mean game score of the first and the last three trials of each exergame, we visualised how the game scores in each participant progressed over the course of time (Figure 4.5). All participants reached higher game scores for Treasure Hunt than when playing Sky Race (Figure 4.5).

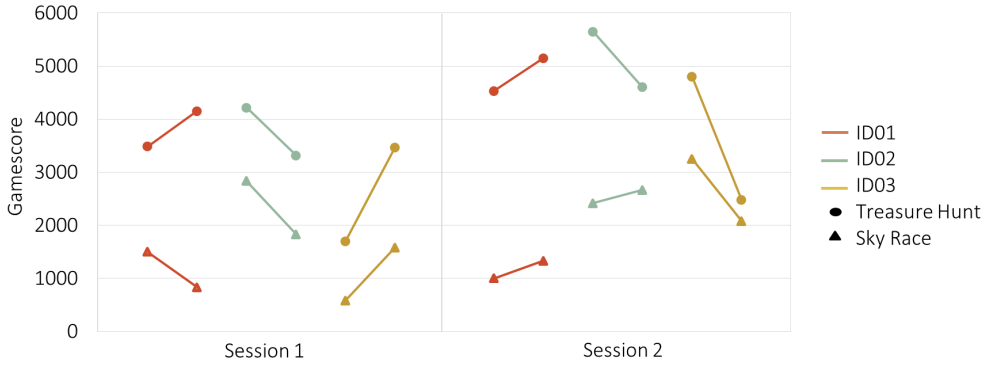


Fig.4.5. Acquisition and retention. Displayed are means of the scores of the first three trials (first data-point of each pair) and of the last three trials (second data-point of each pair). ID01 practiced in random order while ID02 and ID03 practiced in blocked order. Retention is determined as the difference of the last data-point of session one and the first data-point of session two.

Transfer

Obtaining and evaluating Circle and QoM data posed some difficulties. As mentioned earlier, we had a relevant data loss due to the technical issues with ChARMin. For none of the participants, a full dataset was available (Figure 4.6).

All participants scored higher in the BBT after playing the game in session 1 compared to before (ID01 +2 blocks, ID02 + 1 block, ID03 + 4 blocks) (Figure 4.7).

When comparing the performance at delayed transfer with the performance at the end of session 1, ID001 retains his performance, ID02 has a higher (+ 3.5 blocks) score, while ID03 has a lower score (- four blocks) (Figure 4.7). During session 2, only ID01 scored higher from pre to post (+ 0.5 blocks), ID02 performed worse (- 2.5 blocks), and ID03 showed no difference from pre to post.

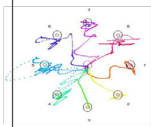
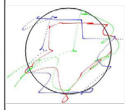
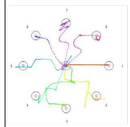
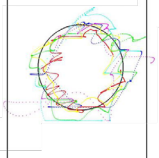
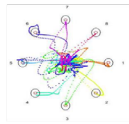
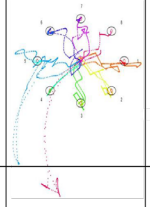
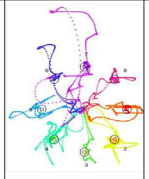
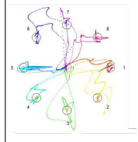
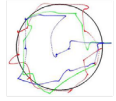
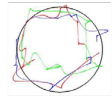
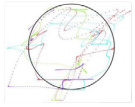
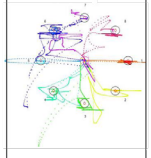
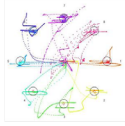
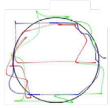
	Day 1 pre		Day 1 post		Day 1 delayed		Day 2 pre		Day 2 post	
ID	QoM	Circle	QoM	Circle	QoM	Circle	QoM	Circle	QoM	Circle
01					No data available	No data available			No data available	No data available
02	No data available	No data available								
03	No data available		No data available						No data available	No data available

Fig. 4.6. Incomplete datasets of robotic assessments. Displayed are the movement paths of the first trial of each time point for the Quality of Movement (QoM) and of all trials of the Circle following (Circle) assessment of each participant.

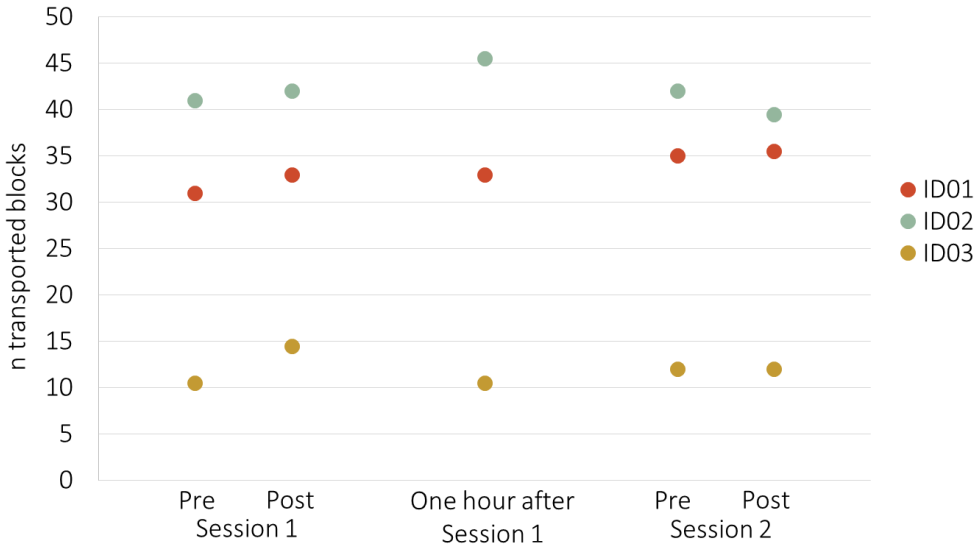


Fig.4.7. Box and Blocks Test. Displayed are the average numbers of transported blocks of two trials for each participant and time-point.

DISCUSSION

We explored the acceptability, the implementation and the practicality of a randomised controlled parallel study design evaluating blocked versus random order in children with brain lesions. In the following, the most important findings of this exploration will be discussed, always in combination with options and ideas for a future study.

The main issue occurring with our exploration were the technical problems we experienced with ChARMin. When the technical issues were inventoried, it became obvious that the majority of the oscillations happened during the appointments with one participant. It is probable, that their movement behaviour provoked the oscillations. The analysis of acceptability showed that compliance was particularly difficult for this same participant. Error messages or oscillations led to delays, breaks and abortions of appointments, which might have had a negative impact on the participant's compliance and motivation. As the engineer was part of the study team, and necessary adaptations were discussed extensively, improvements were implemented soon after these feasibility trials.

From the therapists' point of view, the application of ChARMin was not easy in terms of implementation. Operating the robot and responding to the child in an appropriate manner at the same time, required divided attention from the therapist. The procedure of using ChARMin contained many steps to keep in mind and now and then some of these steps were forgotten (e.g. pushing the safety pedal during

the use). For a future study, it would be of great importance that the involved therapists are trained well and have sufficient experience with the ChARMin so that the operation of the ChARMin (including the correct reaction to error messages) becomes more automated.

An intervention that can be implemented in the clinical context requires some components that we did not think about in detail before. The main question was whether the two exergames chosen for this study would actually generate a contextual interference effect that could be measured in a larger study. Furthermore, we would need to be clear about the motor skill that will be practiced and the outcomes that quantified the skill. We discussed these and other components from the perspective of planning a future study that would enable the evaluation of the contextual interference effect when practicing motor tasks with a robotic device in children with brain lesions.

Motor skill to be practiced

The observation of the contextual interference effect depends on the type of task [1]. In the exergames performed by the participants in this study, different tasks required different kinds of movements. While in Treasure Hunt reaching fish and treasure chests were single, goal-directed start-stop movements, steering a plane or a bird in Sky Race was much more of a continuous movement that required online adjustments. In the planning phase, this was a comprehensible choice from the existing exergames. Yet, what we are looking for is a clinically relevant motor skill that can be practiced with different variations thereof. When looking at human motor development, children start to reach for objects of interest at around the age of 3 months [13]. Reaching and grasping are constituents of functional tasks [14], and this is precisely what is often difficult for children with cerebral palsy and other sensorimotor disorders [15]. Although reaching needs to be combined with functional grasping to become relevant for daily life activities, focusing only on the evaluation of reaching would allow to include also children with more severely affected hand function. Reaching could be the motor skill to practice for a future study. Variations of reaching, implemented in either blocked or random order, would define the two interventions.

Contextual interference is a continuum ranging from low to high. Blocked practice order is located at the lower end, random practice order at the upper end of this scale [1]. Blocked and random practice are therefore providing the highest contrast in terms of the practice order. The indispensable factor of the contextual interference effect is the learning of different variations of a skill and the level of the interference is established by the different characters of the variation [1]. Magill and Hall summarised evidence and concluded that variations of a skill need to be controlled by different motor programmes, i.e. relative timing, sequence of events and/or spatial configurations need to differ across the variations to generate the

typical outcomes
due to contextual interference [1]. A variation of reaching, for example, involving different motor programmes could be attained by performing the reaching movement in different planes. When varying the reaching movement vertically (up and down, left and right on a frontal plane) and reaching horizontally (back and forth, left and right, on a transversal plane), the spatial configurations are varied, thus involving two different motor programs.

Repetitions

During the exergames the children were free to choose the paths on which they wanted to move. Some chose the safe way and steered the airplane over the houses to not crash into any obstacles. Others played more dauntless, chasing the objects that gave points, and trying to avoid the obstacles. The more risky players performed much more movements compared to the careful players. The game behaviour therefore plays an important role in the number of movements. Furthermore, probably there was also a difference in the number of repetitions between the two exergames as they require different motor aspects. The number of repetitions can be looked at differently. First, the number of trials, which are in our case, the repetitions of the exergames. Secondly, there is the number of movements, which in our case, because of the abovementioned different game behaviour, cannot be determined. To generate a contextual interference effect, it is important that the number of repetitions between the different variations (exergames) are equal for both practice groups to make sure that the practice order is the only thing differentiating the practice groups [16].

But how many repetitions are needed to induce motor learning? In adults after stroke, 500 repetitions of a movement led to motor learning effects [17]. During an eight-weeks schedule of robot-assisted therapy consisting of 16 sessions of 640 point-to-point movements each, children with cerebral palsy performed a total of 10240 movements of a game that was similar to our QoM assessment [18]. They showed improvements in clinical measures as well as in robot kinematic metrics [18]. Children with cerebral palsy playing a robot-assisted exergame performed approximately 3000 repetitions of a goal-directed movement which resulted in changes due to motor learning [19]. Depending on the exergames applied in a future study, the number of repetitions should vary somewhere between 500 and 3000 to induce motor learning. Hence, if reaching was practiced as intensively as possible, as many reaching movements as possible should be implemented within one trial of an exergame. Yet, as it is well known, the number of repetitions is crucial in motor learning, but so is motivation [20].

Motivation

The main reasons for aborting the practice sessions during this experiment were the lack of motivation. It has been stated that play motivates repetitions like no other activity [21]. Additionally to the potential influence of the technical problems on the compliance, the exergames we used just seemed not to be motivating enough for the participants. Hence, more motivating exergames and variations are required. Including the participants in choices of exergame styles, scenarios or avatars during playing would be one option to increase the interest and attendance. Indisputably, the number of repetitions and the motivation are closely linked. Optimally, high motivation leads to a higher number of repetitions.

Outcome measures

Acquisition and retention

For this exploratory experiment, we chose the exergame scores as parameters to assess acquisition and retention. As a parameter, the game score has proven to be a poor choice. First, Treasure Hunt generally released higher scores than Sky Race (Figure 4.5). This was not a sign of better performance but because the scores are calculated completely different in the two games. The same parameter should be measured during each repetition, independent on the variation of the skill. Second, only one data point was obtained per exergame trial, which does not sufficiently reflect the performance of the whole trial. Hence, if reaching would be practiced, the parameter of interest for acquisition and retention should be the same, independent of the plane in which reaching is practiced. Each movement should provide a data point to enable measurements of the performance during the whole trial.

Transfer

The outcome measures (i.e. QoM and Circle) we chose to assess the transfer in this exploratory feasibility study, were not suitable. It seemed to be difficult for some participants to move to the upper quadrants of the reference circle and reach the upper targets during the Circle and the QoM assessment, respectively. We observed this already during the reliability study [5]. It certainly required strength and endurance to lift the arm and stay on the target. For some more severely affected children, this could even mean to reach a limit of their strength. Also, the will to summon up the required strength should not be underestimated. A lack of this will or the motivation to go to physical limits could also have been a reason for not reaching the upper quadrants of the screen. As it can be seen in Figure 4.6, also the lower quadrants of the circle to be followed were omitted. Such ‘short cuts’ might rather be a sign of lack of motivation to follow the instructions of the assessments. The ChARMin assessments should be developed further and made more motivating for the children. For example, they could be implemented in exergames. Technical problems should be solved so that the assessments could be applied successfully in a future study. Thus, additional engineering and programming work was required.

Choosing a primary outcome

Although it is possible to choose several primary outcomes, the clinical interpretation is challenging when there are several outcomes reporting conflicting results [22]. In a larger study, choosing different outcome parameters as a primary outcome measure without correcting for multiple comparisons would be fishing for p-values [23]. The use of multiple primary outcomes induces problems related to the multiciplity of analyses and are therefore not recommended [24]. Defining a single primary outcome a priori is therefore decisive. The primary outcome is defined as the in advance specified outcome with the greatest importance to the relevant stakeholders (e.g. patients and/or clinicians) [24]. A primary outcome in our case would mean one parameter of a specific assessment obtained during a specific phase of motor learning (acquisition, retention, or transfer) and assessed at a predefined time point (e.g. during practice of session 1 or session 2, or immediately after, or one hour after, or at follow up). Furthermore, from a clinical point of view, it seems necessary to be able to transfer learned skills in activities relevant in daily life. Skills that work fine in therapy mean nothing if they cannot be generalised to outside the therapy room. Therefore, choosing some kind of transfer effect as primary outcome seems reasonable. Depending on the skill, which is practiced during the intervention, the transfer should contain a similar movement characteristic but should be implemented in a movement or action, which is more related to a daily action. Assessments that measure movement quality capture small changes and are useful to measure the effectiveness of interventions. [25]. Therefore, an outcome reflecting movement quality might be called for. One of the subscales of MA2 could be an option. Secondary outcomes could then cover the acquisition and retention phases.

All three, transfer, retention and acquisition should measure the same construct, namely the one which is practiced. If for example movement fluency is the chosen construct, measures and parameters of transfer, acquisition and retention should be a proxy for this.

CONCLUSION AND OUTLOOK

Conducting a study by using the design of this experiment is not feasible. We performed this experiment with available resources (e.g. the exergames that were provided) and fitted the rationale around it. The next approach should start with the rationale and then organise the resources accordingly (e.g. new exergames, programmed to fit our requirements).

To assess the influence of difference levels of contextual interference, namely blocked versus random practice order, we will have to consider the abovementioned aspects. Consequently, we will programme new and more motivating exergames

for the ChARMin robot. We aim for exergames requiring reaching movements to be played in the frontal and transversal plane (i.e. to generate a variation of reaching, which is based on different motor programmes). We will decide on a primary outcome regarding the phase of motor learning (most probably transfer), a measure (e.g. the subscale "fluency" of the MA2) and a time point (e.g. immediately after the practice phase). For the secondary outcomes we will chose a parameter of interest, which is recorded during playing the exergames and which reflects the outcome for transfer. We will define a study protocol that fits into the rehabilitation schedule and is feasible for the participants. It is reasonable to plan and perform a pilot study as there are still many uncertainties needed to be solved before starting with a larger randomised controlled trial.

ACKNOWLEDGEMENTS

A very special thank you goes to the children participating in this feasibility experiment and their parents.

AUTHOR'S CONTRIBUTION

Conceptualization: Judith V Graser, Hubertus JA van Hedel
Formal analysis: Judith V Graser
Funding acquisition: Hubertus JA van Hedel
Investigation: Judith V Graser
Methodology: Judith V Graser, Hubertus JA van Hedel, Caroline HG Bastiaenen, Rob RA de Bie
Project administration: Judith V Graser
Resources: Hubertus JA van Hedel
Validation: Judith V Graser, Hubertus JA van Hedel
Visualization: Judith V Graser
Writing – original draft: Judith V Graser
Writing – review & editing: Judith V Graser, Hubertus JA van Hedel, Caroline HG Bastiaenen

REFERENCES

1. Magill RA, Hall KG. A review of the contextual interference effect in motor skill acquisition. Hum Mov Sci. 1990;9:241–89.
2. Graser JV, Bastiaenen CH, van Hedel HJA. The role of the practice order: a systematic review about contextual interference in children. PLOS One. 2019;14.
3. Keller U, Van Hedel HJA, Klamroth-Marganska V, Riener R. ChARMin: The first actuated exoskeleton robot for pediatric arm rehabilitation. IEEE/ASME Trans Mechatronics. 2016;21:2201–13.
4. Keller U, Herold-Nadig M, Kollmar A. ChARMin Benutzerhandbuch, Version 1.0, 12.01.2015. Sensory-Motor Systems Lab, Federal Institure of Technology, Zurich.
5. Graser JV, Prospero L, Liesch M, Keller U, van Hedel HJ. Reliability of and practical lessons learned from robotic upper limb assessments in children undergoing neurorehabilitation. Submitted.
6. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Ohrvall A-M, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: Scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006;48:549–54.
7. Randall M, Johnson L, Reddihough D. The Melbourne Assessment 2. https://www.rch.org.au/melbourneassessment/. Accessed 29 Mar 2019.
8. Wang T-N, Liang K-J, Liu Y-C, Shieh J-Y, Chen H-L. Psychometric and clinimetric properties of the Melbourne Assessment 2 in children with cerebral palsy. Arch Phys Med Rehabil. 2017;98:1836–41.
9. Ritter N, Kilinc E, Navruz B, Bae Y. Test Review: L. Brown, R. J. Sherbenou, & S. K. Johnsen 'Test of Nonverbal Intelligence-4' (Toni-4). Austin, TX--PRO-ED, 2010. J Psychoeduc Assess. 2011;29:484–8.
10. Mathiowetz V, Federman S, Wiemer D. Box and Block Test of manual dexterity: Norms for 6-19 year olds. Can J Occup Ther. 1985;52:241–5.
11. Jongbloed-Pereboom M, Nijhuis- van der Sanden MWG, Steenbergen B. Norm scores of the Box and Block Test for children ages 3–10 years. Am J Occup Ther. 2013;67:312–8.
12. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. Am J Occup Ther. 1985;39:386–91.
13. Shadmehr R, Mussa-Ivaldi FA. Adaptive representation of dynamics during learning of a motor task. J Neurosci. 1994;14:3208–24.
14. Nicholson JH, Morton RE, Attfield S, Rennie D. Assessment of upper-limb function and movement in children with cerebral palsy wearing lycra garments. Dev Med

Child Neurol. 2001;43:384–91.

15. Gordon AM, Duff S V. Relation between clinical measures and fine manipulative control in children with hemiplegic cerebral palsy. Dev Med Child Neurol. 1999;41:586–91.
16. Lee TD, Simon DA. Contextual interference. In: Williams AM, Hodges NJ, editors. Skill acquisition in sport: Research, theory and practice. London and New York: Routledge; 2004. p. 29–44.
17. Boyd L, Winstein CJ. Explicit information interferes with implicit motor learning of both continuous and discrete movement tasks after stroke. J Neurol Phys Ther. 2006;30:46–57.
18. Krebs HI, Fasoli SE, Dipietro L, Fragala-Pinkham M, Hughes R, Stein J, et al. Motor learning characterizes habilitation of children with hemiplegic cerebral palsy. Neurorehabil Neural Repair. 2012;26:855–60.
19. Keller JW, van Hedel HJ. Weight-supported training of the upper extremity in children with cerebral palsy: A motor learning study. J Neuroeng Rehabil. 2017;14.
20. Shmuelof L, Krakauer JW, Mazzoni P. How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. J Neurophysiol. 2012;108:578–94.
21. Scarfe N V. Play is education. Child Educ. 1962;39:117–21.
22. Tyler KM, Normand S-LT, Horton NJ. The use and abuse of multiple outcomes in randomized controlled depression trials. Contemp Clin Trials. 2011;32:299–304.
23. Andrade C. The primary outcome measure and its importance in clinical trials. J Clin Psychiatry. 2015;76:1320–3.
24. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Br Med J. 2010;340:1–28.
25. Demers M, Levin MF. Do activity level outcome measures commonly used in neurological practice assess upper-limb movement quality? Neurorehabil Neural Repair. 2017;31:623–37.

CHAPTER 5

CONTEXTUAL INTERFERENCE IN CHILDREN WITH BRAIN LESIONS: PROTOCOL OF A PILOT STUDY INVESTIGATING BLOCKED VS. RANDOM PRACTICE ORDER OF AN UPPER LIMB ROBOTIC EXERGAME

Judith V Graser
Caroline HG Bastiaenen
Urs Keller
Hubertus JA van Hedel

PAFS 2020; 6(156)

ABSTRACT

Background: If adults practice several motor tasks together, random practice leads to better transfer and retention compared to blocked practice. Knowledge about this contextual interference (CI) effect could be valuable to improve neurorehabilitation of children. We present the protocol of a randomised controlled pilot study investigating the feasibility of blocked practice vs. random practice of robot-assisted upper limb reaching in children with brain lesions undergoing neurorehabilitation.

Methods: Children with affected upper limb function due to congenital or acquired brain lesions undergoing neurorehabilitation will be recruited for a randomised controlled pilot study with a 3-week procedure. In the control week (1), two assessment blocks (robot-assisted reaching tasks, Melbourne assessment 2, subscale fluency), 2 days apart, take place. In the practice week (2), participants are randomly allocated to blocked practice or random practice and perform 480 reaching and backward movements in the horizontal and vertical plane using exergaming with an exoskeleton robot per day during three consecutive days. Assessments are performed before, directly after and 1 day after the practice sessions. In the follow-up week (3), participants perform the assessments 1 week after the final practice session. The primary outcome is the immediate transfer of the Melbourne Assessment 2, subscale fluency. Secondary outcomes are the immediate retention, 1-day and 1-week delayed transfer and retention and acquisition during the practice sessions. We will evaluate the feasibility of the inclusion criteria, the recruitment rate, the scheduling procedure, the randomisation procedure, the procedure for the participants, the handling of the robot, the handling of the amount of data, the choice of the outcome measures and the influence of other therapies. Furthermore, we will perform a power calculation using the data to estimate the sample size for the main trial.

Discussion: The protocol of the pilot study is a first step towards a future main randomised controlled trial. This low risk pilot study might induce some benefits for the participants. However, we need to place its results into perspective, especially concerning the generalisability, as it remains questionable whether improving reaching constrained within a robotic device will ameliorate daily life reaching tasks.

Trial registration: ClinicalTrials.gov Identifier: NCT02443857

BACKGROUND

The population of children undergoing neurorehabilitation is as heterogeneous as the causes for brain dysfunction. Differentiating between congenital and acquired brain lesions is common. Children with congenital brain lesions are diagnosed most frequently with cerebral palsy. The prevalence of cerebral palsy seems to decrease in Europe (from 1.90 to 1.77 per 1000 live births between 1980 and 2003); it is still the most common cause for severe motor impairments in children [1]. Acquired brain injuries in children are less common. In Norway, for example, the incidence for moderate traumatic brain injury amounts to 0.024 per 1000 children and for severe traumatic brain injury 0.025 per 1000 children [2]. For stroke, the worldwide incidence ranges between 0.013 and 0.13 per 1000 children [3]. Both, congenital and acquired brain lesions interfere with the normal development of the brain, which causes impairments in sensorymotor and cognitive functions and limitations in activities that may significantly impact quality of life [4]. Hence, these children should be referred to paediatric neurorehabilitation.

Nowadays, most neuro-rehabilitation programmes, including those for paediatric neurorehabilitation, are based on motor learning principles [5–7]. The general goal of motor learning is to attain relatively permanent changes in movement skills by practice and experience [8]. It has been suggested that the performance during practice (i.e. acquisition), retention (i.e. the preservation of the learned skill for a certain period) and transfer (i.e. when transferring the learned skill to another task) needs to be distinguished [9]. While retention is sustainability of performance after a practice phase, transfer reflects the effect of the practice on other, yet untrained, tasks [9]. Transfer is especially important since it allows generalising improved motor functions or capacities to daily life performance.

If several tasks have to be learned and are practiced within the same therapy session, they can interfere with each other, which could affect the outcome. In this so called contextual interference effect [10], the practice order is an important factor that determines the strength of this effect. If one task is practiced several times before switching to the next task (i.e. practicing in a blocked order), contextual interference is low [10]. If different tasks are practiced in a random order, interference is high [10]. Contextual interference has mainly been investigated in healthy adults where it was shown that high contextual interference leads to worse performance in acquisition but better performance in retention and transfer, while low contextual interference leads to contrary results [11, 12]. Two main hypotheses explaining these findings have been discussed in the literature. On the one hand, the elaborative-processing hypothesis states that compared to blocked practice, the learning process during random practice is based on a more profound elaboration of the tasks due to the comparisons between and within the trials [13]. This could lead to a more

comprehensive memory trace, which is easier to retrieve [14]. The forgetting-reconstruction hypothesis, on the other hand, is based on a strengthened memory consolidation occurring during random practice [15]. The underlying explanation is that during random order practice, the learner switches between different tasks all the time and forgets the established action plan of the prior task when a subsequent new task is performed [14]. It seems comprehensible that (a) a deeper elaborated and more robust memory representation could result in better retention and transfer and (b) acquiring a motor skill under blocked practice might be easier because of less disturbance, leading to a better momentary performance at acquisition compared to random practice.

It has been stated that the contextual interference effect is larger when the tasks involve different motor programmes [11, 16]. A motor programme can be understood as a memory for a movement class rather than for an action or a movement solely [17, 18]. Actions that have the same invariant aspects in common (e.g. spatial and topographical characteristics of the action [19], relative force and timing and sequences involved in the action [11]) are controlled by the same motor programme [11, 16]. Therefore, applying variations of the invariant features would increase the contextual interference effect. However, as contextual interference has mainly been investigated in healthy adults, evidence is lacking for children with congenital or acquired brain lesions [20].

Adhering to the recommendation of providing intensive and repetitive training to restore motor function in neurorehabilitation [21] is challenging, especially when working with children. To keep them engaged during highly repetitive therapy sessions, practice has to be variable. To induce variability, several tasks or variations of a task are practiced usually within one therapy session. Since we aim to provide efficient therapies to obtain optimal outcomes, we consider it relevant to improve our understanding of the influence of contextual interference on the therapeutic outcome of paediatric neurorehabilitative interventions.

In recent years, rehabilitation technologies have been applied increasingly also in paediatric neurorehabilitation [22]. Main advantages are the standardisation of training protocols and assessments, high number of repetitions, and above all, additional motivation due to exergames [22]. While the definition of 'exergame' is debated, in this study, we use the common definition 'videogames that require physical activity in order to play' p. 10 [23]. To our knowledge, two studies used new technologies to investigate the influence of the practice order. The first study investigated motor learning in children with developmental coordination disorder and typically developing children. The children practiced with the Nintendo® Wii Fit video game during 10 20-min sessions in a variable group (a self-selected choice out of 10 games) and a repetitive group (the same game throughout the whole session) [24]. In the second study, children with cerebral palsy and typically developing

children practiced a computer-maze in random order (30 trials, 5 different mazes in random order) and in constant order (30 trials of 1 maze) [25]. While both practice groups improved their game performance similarly in the first study [24], the random order group showed a reduction in movement time needed to complete the maze tasks in retention and transfer in the second study [25]. However, as both studies did not include a true blocked group, we designed a study protocol to investigate blocked vs. random order. As there is not much knowledge in the field of contextual interference with robotic exergames in paediatric neurorehabilitation, we planned a pilot study to evaluate the feasibility of a future main study. The objective of this paper was to present the protocol of this pilot study evaluating the feasibility of a randomised controlled, single-blinded study about contextual interference in robot-assisted upper limb training in children with congenital or acquired brain lesion and affected upper limb function. To assess the feasibility, we aimed to address the following questions:

1. Are the chosen inclusion criteria specific enough to result in a sample of participants, which is suitable for this pilot study?
2. Is the recruitment rate feasible?
3. Is the scheduling procedure feasible?
4. Is the randomisation procedure feasible?
5. Is the whole procedure feasible for the participants?
6. Is the handling of the robot feasible?
7. Is the handling of the large amount of data feasible?
8. Are the outcome measures responsive and sensitive enough within this setting?
9. Is there a confounding influence of parallel therapies within the rehabilitation setting?
10. Is it feasible to conduct the main trial with respect to the needed sample size calculated from the data obtained for the primary motor learning outcome?

METHODS

Design

This pilot study protocol describes a randomised controlled single-blinded two-arm intervention study comparing several outcome time points between two groups of children with neuro-motor upper limb impairment who will practice two variations of reaching movement tasks with robot-assisted exergames (one group in blocked, the other in random order) with a predefined primary outcome and a follow-up period of 1 week. We will include a follow-up period of 1 week to match the

rather short intervention period; 1 week will also be feasible to plan within a rehabilitation stay.

Setting

The study will take place at the Swiss Children's Rehab of the University Children's Hospital Zurich in Affoltern am Albis, Switzerland. All measurements and practice sessions will be performed during a multidisciplinary inpatient rehabilitation stay.

Ethical considerations and reporting

This study has been approved as part of the ChARMin project (sub-project 5: Motor learning) by the Ethics Committee of the Canton Zurich (BASEC-Nr. PB_2016-02450) and the Swiss Agency for Therapeutic Products (Swissmedic reference number: 2015-MD-0009).

According to the Ethics Committee's guidelines, all the participants will give their verbal consent, children of 15 years and older and all the legal representatives will provide informed written consent.

This study protocol was established according to the guide on reporting protocols of pilot and feasibility trials [26], the Consolidated Standards of Reporting Trials (CONSORT) statement extension for randomised pilot and feasibility trials [27] and the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) [28]. See Additional file 1 for the SPIRIT checklist and Figure 5.1 for the SPIRIT Figure.

Participants

The majority of children admitted to our rehabilitation centre has a congenital brain lesion (cerebral palsy) but we treat also children with acquired brain lesions (after stroke or traumatic brain injury), spinal cord injury, genetic syndromes, etc. The reasons for admitting a patient vary. Most children with cerebral palsy come to the centrefor a shorter period (4 to 6 weeks) of intensive therapy or after orthopaedic surgery, particularly of the lower limbs, with consecutive rehabilitation. Patients who experienced an acquired brain lesion, such as a stroke or traumatic injury, are admitted as soon as they are stable and can leave the intensive care unit. Depending on the recovery, some of these patients might also be re-admitted for rehabilitation at a later stage. Currently, the average length of stay is around 40 days but varies from a couple of weeks to more than a year (for children with acquired brain lesions). According to their rehabilitation goals, the children undergo physiotherapy, occupational therapy, sports therapy, speech and language therapy, neuropsychology, hippotherapy, and robot-assisted therapy. For this pilot study, we will recruit a sample representing patients with congenital or acquired brain lesions, reflecting the majority of our inpatient population.

		Study period								
		Enrolment	**Allocation**	**Control period**		**Practice period**			**Follow-up**	
Timepoint		Before day 1	Before day 8	Day 1	Day 3	Day 8	Day 9	Day 10	Day 11	Day 17
Enrolment	Eligibility screen	X								
	Informed consent	X								
	Obtaining parameters for minimisation	X								
	Allocation		X**							
Assessments	Test of Nonverbal Intelligence	X*								
	Melbourne Assessment 2			X	X	X		X	X	X
	Game test			X	X	X		X	X	X
Interventions	Blocked practice					X	X	X		
	Random practice					X	X	X		

Fig. 5.1. Standard Protocol Items Recommendation for Interventional Trials (SPIRIT) figure. Actions and appointments throughout the whole study. *The Test Of Nonverbal Intelligence is conducted before the first practice session (outcome needed for randomisation by minimisation). **Allocation is done when all the parameters for minimisation are obtained; at latest on the day before the first practice session.

Inclusion and exclusion criteria

Included will be children with either congenital or acquired brain lesions. Those with an acquired brain lesion should be in the subacute, i.e. more than 3 months, or chronic stage. Further inclusion criteria will be uni- or bilaterally affected upper limb function with spasticity,dyskinesia or mixed conditions. Additional inclusion criteriawill be age 5 to 18 years, the ability to sit upright for approximately 60 min without lateral trunk support and a Manual Ability Classification System (MACS) level between I and IV (MACS level I: handles objects easily and successfully; level II: handles most objects but with somewhat reduced quality and/or speed of achievement; level III: handles objects with difficulty, needs help to prepare and/or modify activities; level IV: handles a limited selection of easily managed objects in adapted situations [29]).

Moreover, the children will need to be able to understand and follow test instructions, be compliant for the whole study procedure, can communicate pain or discomfort and see a computer screen at approximately 1m in front of him/her. Excluded will be children with upper limb surgery or Botox injections during the past 6 months, and upper limb skin lesions.

We decided not to include outpatients since we learned from prior experiences that it is difficult for parents to organise their children's attendance for participating in such an extensive study procedure.

Target sample size

Since the recruitment rate will be one of our feasibility criteria, we will recruit for 1 year and calculate the recruitment rate, taking into account the number of eligible and recruited participants, and complete datasets. Yet, the CONSORT 2010 statement extension to randomised pilot and feasibility trials recommends to give some rationale for the target sample size [27]. Based on our center's numbers from past years and taking into account the inclusion criteria, we anticipate to recruit 20 participants (10 per group).

Recruitment

We will inform the children who are admitted to our centre matching the inclusion criteria and their parents/legal representatives about the study. If they provide their written informed consent, the children will be included and the appointments scheduled.

Group allocation

Participants will be allocated to a blocked or random practice group using randomisation by minimisation. This method enables balancing several prognostic factors in small samples [30]. We will impute parameters that potentially influence motor learning. To reduce the risk of selection bias, these factors should be balanced between the groups, even in the case of uneven distribution.

We will use the following minimisation parameters: age (preschool age: 5–6 years, primary school age: 7–12 years, secondary school age and older: 13–18 years); gender (female, male); diagnose (congenital, acquired); manual ability (MACS level I, II, III or IV); and cognitive ability according to the Test of Nonverbal Intelligence–Fourth edition (TONI-4), which evaluates abstract reasoning and problem solving (Index Score < 70: very poor, 70–79: poor, 80–89: below average, 90–110: average, 111–120: above average, 121–130: superior and > 130: very superior) [31]. The minimisation parameters are equally weighted. A study nurse will receive the minimisation parameters of each new participant by one of two researchers involved in recruitment. She will enter the parameters in a custom-written Matlab programme, which will allocate the participant to either the blocked or random order group. The study nurse will be unaware of this definition and the purpose of the study. The study nurse will enter the allocation in a file, which will allow the researcher to schedule the appointments and perform the practice and measurement sessions. The study nurse will not be involved otherwise in this trial. In case the study nurse is unavailable, we will assign another independent person who will perform this procedure.

Equipment, intervention, outcome measures and study procedure

Robot

We selected the ChARMin (Children's Arm Rehabilitation Mechatronic Interface) device that was developed in a collaboration between the Sensory-Motor Systems Lab (Federal Institute of Technology Zurich, Switzerland) and the Swiss Children's Rehab [32]. ChARMin (Figure 5.2) is an actuated exoskeleton robot, which can actively support joint movements of the shoulder, elbow, forearm and wrist for the left or the right arm. To support the patient's movements, the motors compensate the weight and friction of the robot arm and can provide guidance force. Therapists can adjust the guidance from a non-supporting mode, where the robot transparently follows the patient's movement, to a fully supported mode. This allows adapting the physical support of ChARMin to each patient during therapies. However, for the planned pilot study, participants will not receive physical guidance or support of the device; the motors will only compensate for the weight and friction of the robot.

ChARMin includes a small and large distal module to fit the anthropometrics from about 5-year-old children to adolescents. Interfaced with different exergames, it provides and facilitates playful training of arm and hand functions, especially for children with more severe upper limb impairments. During the execution of the exergames, different sensors and game parameters are recorded and processed to extract the assessment measures (e.g. number of velocity peaks, precision on the target or reaction time).

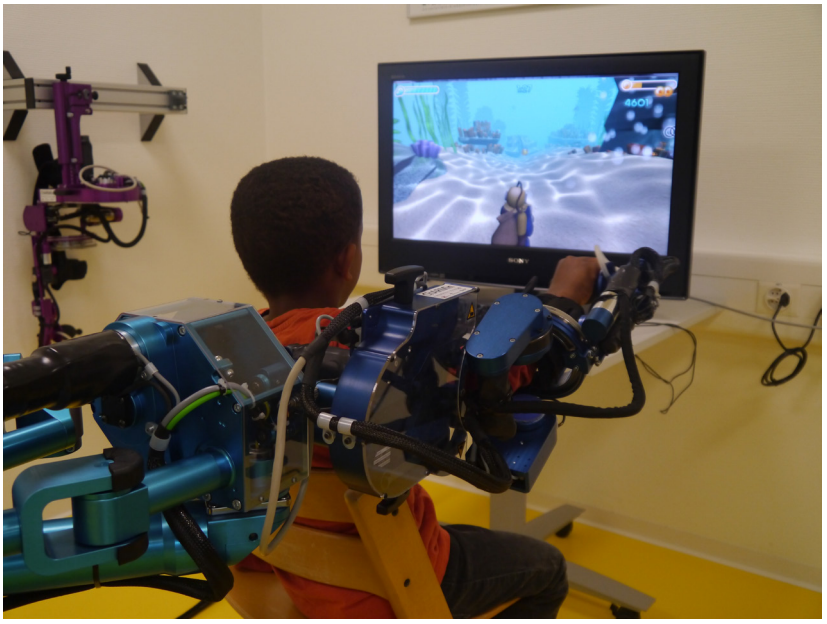


Fig. 5.2. The ChARMin robot. The ChARMin exoskeleton: A young boy practicing with ChARMin's small distal module.

Intervention

We selected an exergame for the pilot study that is based on the Quality of Movement assessment provided by ChARMin [33]. The participants will need to perform goaldirected reaching movements. The participants will have to steer an avatar, which represents the position of the participant's hand, on the screen by moving the upper extremity with the attached exoskeleton towards one of eight targets appearing in random order radially around the centre. The children will be instructed to reach the targets by the most direct path and as fast as possible. The avatar will need to remain on the target for 2 s. After that, the target will disappear. Then, the centre object will reappear, and the avatar will have to be moved back to the centre object to remain there for another 2 s before the next target will appear. If the participant will not reach the target within 10 s, it will disappear, and the centre object will reappear.

Since interference levels must be appropriate to produce a contrast between random and blocked schedules [34], we developed two versions of the exergame; one can be played in the transversal plane, the other in the frontal plane (i.e. the two versions of the exergames have different spatial characteristics, which should require different motor programmes) [11].

Choosing different varying characteristics would be an option, and has already been evaluated [35]. Including variations based on different motor programmes has been suggested to create a more difficult learning situation [11]. As we did not want to increase the learning difficulty too much for the participants, we decided to vary the tasks only in one parameter. In the transversal plane version, the exergame will be displayed as a horizontal plane on the screen. The participant will move inside a haptic wall, i.e. the movement will be restricted mechanically to a horizontal plane, which will be located 10 cm below the shoulder joint for the small and the large distal module (Figure 5.3 (A1 and A2)). Mainly horizontal shoulder adduction and abduction, and elbow flexionand extension will be required to play the horizontal version of the exergame. In the frontal plane version, the exergame will be displayed vertically on the screen. Here, the participant will move inside a haptic wall located 30 cm in front of the shoulder joint for the small distal module and 35 cm in front of the shoulder joint for the large distal module (Figure 5.3 (B1 and B2)). Shoulder flexion and extension, (horizontal) abduction and adduction, shoulder internal- and external rotation and some elbow flexion and extension will be required to play the vertical version of the exergame.

To motivate the participants to perform a large number of repetitions actively, we developed different avatartarget-exergame-scenarios (Figure 5.4). Prior to the exergame, the participant will select his/her own avatar to improve compliance. For the transversal plane exergame version, the avatar will either be a unicorn who will be eating cupcakes (targets) or a snail eating apples. For the frontal plane exergame,

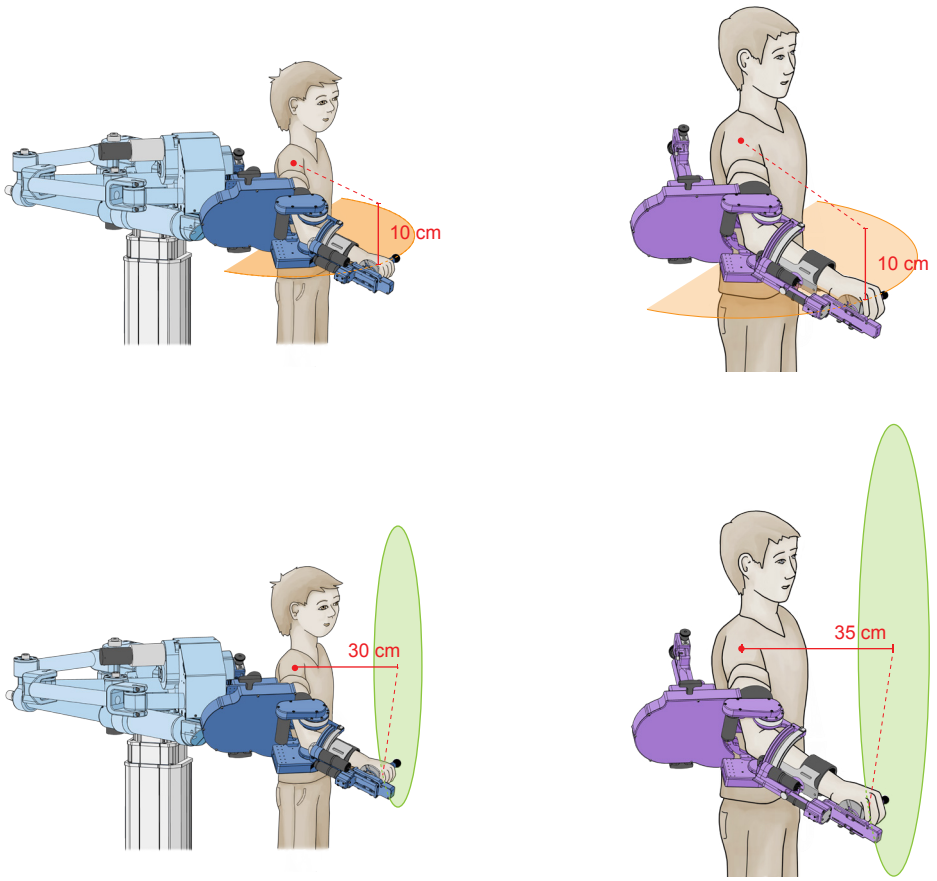


Fig. 5.3. Restrictions of movement inside a haptic wall on the horizontal (A) and the vertical plane (B). Restrictions of movement inside a haptic wall on the horizontal (A) and the vertical plane (B). The movement area during the horizontal plane exergame with the small distal module (A1) and the large distal module (A2). The haptic wall is located 10 cm below the shoulder joint. The movement area on the vertical plane is located 30 cm in front of the shoulder joint for the small distal module (B1) and 35 cm in front of the shoulder joint for the large distal module (B2).

the avatar will be an UFO landing on planets or a submarine collecting fish. As previously mentioned, we will instruct the participants to 'move the unicorn/snail/ UFO/submarine to the cupcake/apple/planet/fish as direct and fast as possible'.

Reaching a target will be rewarded with a score displayed in the middle of the screen. The score will be calculated using the time remaining to reach the target (0 to 10 s), and the distance to path ratio (1=direct movement to the target):

$$Score = time\ remaining + \left(11 - \frac{path}{distance}\right)$$

Therefore, the maximum score for each target will be 20. A randomly chosen target out of the eight targets per trial will be a golden target with a value worth five times as much to increase the participant's motivation. The current score of the ongoing trial will be displayed in the left upper corner of the screen. The total score will be displayed after each trial and will not exceed 240 points.

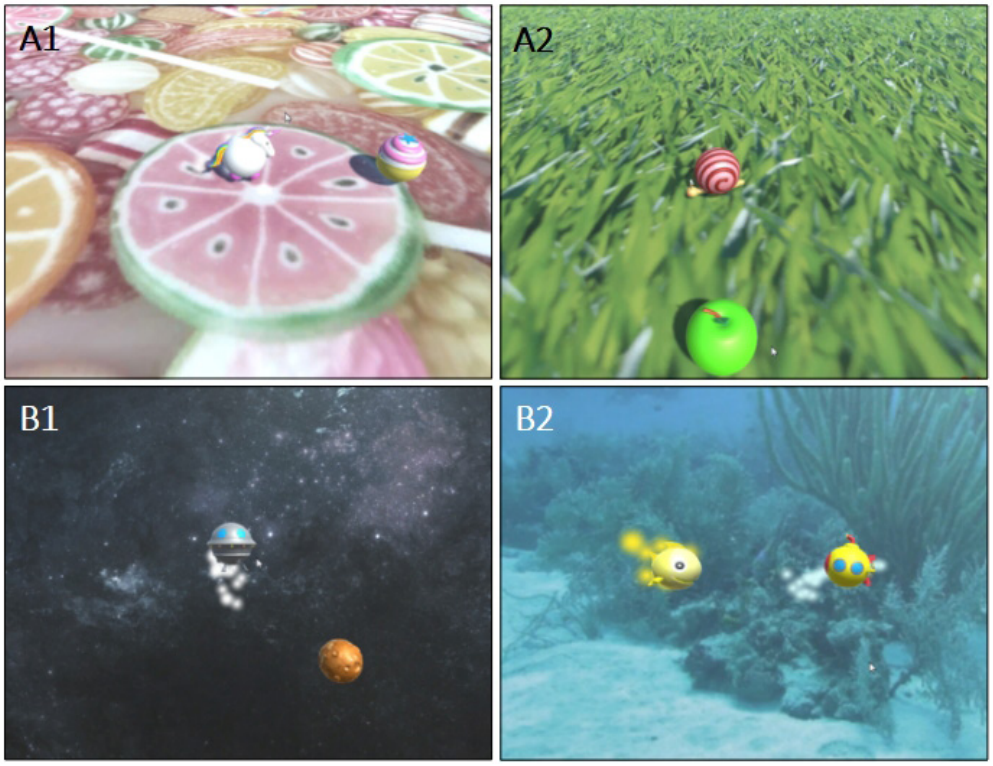


Fig. 5.4. ChARMin exergames. Upper row: transversal version of the ChARMin exergames (A1: avatar unicorn, target cupcake; A2: avatar snail, target apple). Lower row: frontal version of the ChARMin exergames (B1: avatar UFO, target planet; B2: avatar submarine, target fish, in this picture the gold target is displayed).

The blocked group will play 15 trials of one version (i.e. in either the horizontal or vertical plane) followed by 15 trials of the other version. The order of the two blocks will be randomised and the same for all the practice sessions for the same participant. The random group will play the 30 trials of the two versions in a pseudorandom order. The only restriction will be that the same version cannot be repeated more than twice in a row.

To generate observable changes due to motor learning, a large number of repetitions (i.e. hundreds of daily repetitions for upper extremity movements) will be required [36, 37]. Results from a recent study proposed changes in performance scores during

three sessions, each lasting 20–25 min of robot-assisted arm training, and showed that movements were repeated more than 3000 times [38]. Yet, it is challenging to define the number of repetitions sufficient to induce motor learning, while maintaining an acceptable length of sessions to keep the burden for the participants at a minimum and their motivation at a maximal level. In our study, participants will perform 16 reaching movements within one trial of exergames. Over the three sessions (30 trials per session), this would lead to 1440 repetitions of the reaching movement. We consider that this number will induce change due to motor learning. On the other hand, we expect a maximum session duration of 120 min, which seems acceptable.

When switching between the two exergame versions, the therapist will not need to adjust the hardware settings, but she will need to choose the appropriate version of the exergame on the interface. In order to have the same conditions for all participants, we will renounce using the support modalities of the robot. While children with unilaterally affected upper limb function will practice and perform all the assessments with their more affected side, children with bilaterally affected upper limb function will use the arm which they subjectively use more in daily life.

Outcome measures

Outcomes of the main study

Primary outcome - immediate transfer of the practiced skill: We chose the Melbourne Assessment 2 subscale fluency ($MA2_{fluency}$) as the primary outcome measure for the immediate transfer. The MA2 consists of the four subscales movement range, accuracy, dexterity and fluency [39] and is an adequate measurement for research and clinical use in children with cerebral palsy [40]. The intraclass correlation coefficient of the test-retest reliability of the $MA2_{fluency}$ is 0.96 (95% confidence interval 0.90–0.99), the minimal detectable change is 2 points and the minimal clinically important difference is at least 3 points [40]. Concerning the concurrent validity, Pearson's correlation coefficients between the $MA2_{fluency}$ and other tests amounted to 0.67 (Bruininks-Oseretsky Test of Motor Proficiency, Second Edition, subtest 3, manual dexterity), 0.76 (Box and Blocks Test) and 0.40 (Paediatric Motor Activity Log–Revised, quality of movement) [40]. The responsiveness of the $MA2_{fluency}$ is high, which is indicated by significant change scores and a high standardised response mean value of 1.84 [40].

From a clinical point of view and in line with literature (e.g. [9]), we consider the transfer to be an important aspect of learning. We hypothesise that $MA2_{fluency}$ will improve when practicing the reaching movements. During the exergames, the participant will be instructed to reach the target by the most direct path, i.e. as fast as possible. This is in line with the $MA2_{fluency}$ scoring criteria, which include jerkiness, tremor and/or reduced speed of movement [41].

Since we considered the $MA2_{fluency}$ to be more sensitive to change instantly after the

practice sessions due to the relatively short practice period and to avoid a potential loss to follow-up at later time points, we chose the immediate transfer as the time point of interest for the primary outcome.

One of three trained occupational therapists will score the $MA2_{fluency}$ by analysing the video. The rater will be blinded for the participant's group allocation, the time point of assessment and the randomisation code. The sum scores will be calculated and used for further analysis.

Secondary outcomes

- *Delayed transfers:* The delayed transfers (1 day and 1 week after the last practice session) will be determined by the sum score of the $MA2_{fluency}$ at these time points.
- *Immediate and delayed retentions:* To evaluate whether the participants can retain what they have learned, the so-called exergame test will be performed. It consists of one block of three trials of the horizontal version of the exergame and one block of three trials of the vertical version of the exergame. Although the order of the versions will be randomised between the participants, it will remain the same within each participant throughout the study.

  With these exergame tests, we anticipated measuring the construct 'movement fluency'. There is no generally accepted way of quantifying fluency. Several studies have used the absolute number of velocity peaks of a single movement [42, 43]. Yet, as we also wanted to take into account the length of the movement path, we will calculate the parameter 'number of peaks', which is the number of velocity peaks (i.e. when the difference between a local speed minimum and a local speed maximum exceeds a value of 2.5% of the measured maximal speed) normalised to the covered distance of the movement path (nP_{norm}). We will calculate nP_{norm} for each movement (i.e. one value from the centre to the target and one for the return to the centre starting point) that will lead to 16 data points per trial. The parameter nP_{norm} has shown reliable results during the Quality of Movement assessment [33], on which the exergames are based. In a comparable group of 23 participants, the parameter nP_{norm} showed no systematic error between two measurements within 3 to 7 days, an intraclass correlation coefficient of 0.94, a percentage standard error of measurement (divided by the grand mean) of 6.4% and a percentage smallest real difference (divided by the grand mean) of 17.7% (unpublished data).

- *Acquisition:* To observe motor learning during the practice sessions, nP_{norm} will be the parameter of interest too. It will also be recorded 16 times per trial, which will lead to 480 data points per practice session.

Feasibility outcomes

In line with recommendations on reporting protocols of pilot and feasibility trials [26], we will apply the following ten specific feasibility criteria:

1. Are the chosen inclusion criteria specific enough to result in a sample of participants, which is suitable for this pilot study (e.g. are they able to play the exergame)?
2. Is the recruitment rate feasible? Since the main randomised controlled trial would require a large number of participants, it is relevant to know the number of participants that will be recruited during a certain period. The number of children enquired for the study, number of participants recruited, number of complete datasets within 1 year will be recorded and used to calculate recruitment rates.
3. Is the scheduling procedure feasible? The appointments will be planned for 3 weeks for each participant. The person conducting the exergame tests and the $MA2_{fluency}$ always will need to be the same. The same will apply for the person instructing the child during the practice sessions. Absences of personnel (e.g. due to illness, holidays or part-time employment) and the children (e.g. due to illness or short-term planned medical appointments), as well as the availability of the robot and material for the $MA2_{fluency}$ (i.e. whether they are not used for therapy at the time points of the measurements), will need to be considered. Information about whether it is possible to schedule rooms, persons involved and the robot and table for performing assessments will be noted.
4. Is the randomisation procedure feasible? Experiences of the persons involved will be collected and evaluated to answer this question.
5. Is the whole procedure feasible for the participants? The rather time-consuming procedure and the high number of repetitions might be difficult for at least some of the participants because of fatigue or lack of motivation to play the exergames so many times. Both reasons influence motor learning negatively. Therefore, it is important to determine whether the participants included in this study are able to comply the whole study procedure. Information about aborted or shortened sessions or other incidents and their reasons (e.g. pain or discomfort, fatigue) will be taken from the lab journal.
6. Is the handling of the robot feasible? The ChARMin robot has been in use in the clinical and research area since 2015. Since it is a newly developed tool, it is important to evaluate its handling. Any technical issues will be noted and then retrieved from the lab journal.
7. Is the handling of the large amount of data feasible? Information from the assessors and the analyst on the primary outcome will be used.

8. Are the outcome measures responsive and sensitive enough within this setting? Responsiveness analysis of the assessment data will be performed.
9. Is there a confounding influence of parallel therapies within the rehabilitation setting? As the study will be conducted during the inpatient rehabilitation stay, it is important to estimate whether parallel-applied therapies could confound the results. Since an additional control group following usual care would require a larger sample, we chose a design in which each participant acts as its own control and, therefore, we included control measurements during the first week of the study procedure. The changes in the outcomes between week 1 (control week) and week 2 (practice week) will be compared. The kind and number of additional therapies taking place during the different study periods will be recorded to increase the quality of reporting [44].
10. Is it feasible to conduct the main trial with respect to the needed sample size calculated from the data obtained for the primary motor learning outcome? With the data of the primary outcome measure, we will calculate the required sample size for the main trial.

Study procedures

The study procedure is shown in Figure 5.1 (SPIRIT Figure) and Figure 5.5.

Week 1

During the first appointment, the setup of the robot and the adjustments according to the child's anthropometrics will be performed. This takes approximately 15 min. Each participant will then become familiarised with the two versions of the exergame. First, the transversal version 'Unicorn' will be played with manual guidance and verbal support from the therapist. Then, the transversal version 'Snail' will be played with only verbal guidance of the therapist. The same procedure will be followed for the exergame that is played in the frontal plane (i.e. 'Submarine' will be played with manual guidance and verbal support followed by 'UFO', which will be played with only verbal support). Subsequently, an assessment block consisting of the $MA2_{fluency}$ [45] and an exergame test will be performed.

During week 1, also all the parameters for randomisation by minimisation will be obtained from the participant's medical history (e.g. age, gender, diagnosis), from the attending occupational therapist (MACS level), and trained neuropsychologist (TONI-4). Participants will be characterised by age, gender, weight, height, diagnosis (including information about the affected side and specific features such as spasticity, dyskinesia or mixed condition) and the MACS level. After obtaining the minimisation parameters, participants will be allocated to their group.

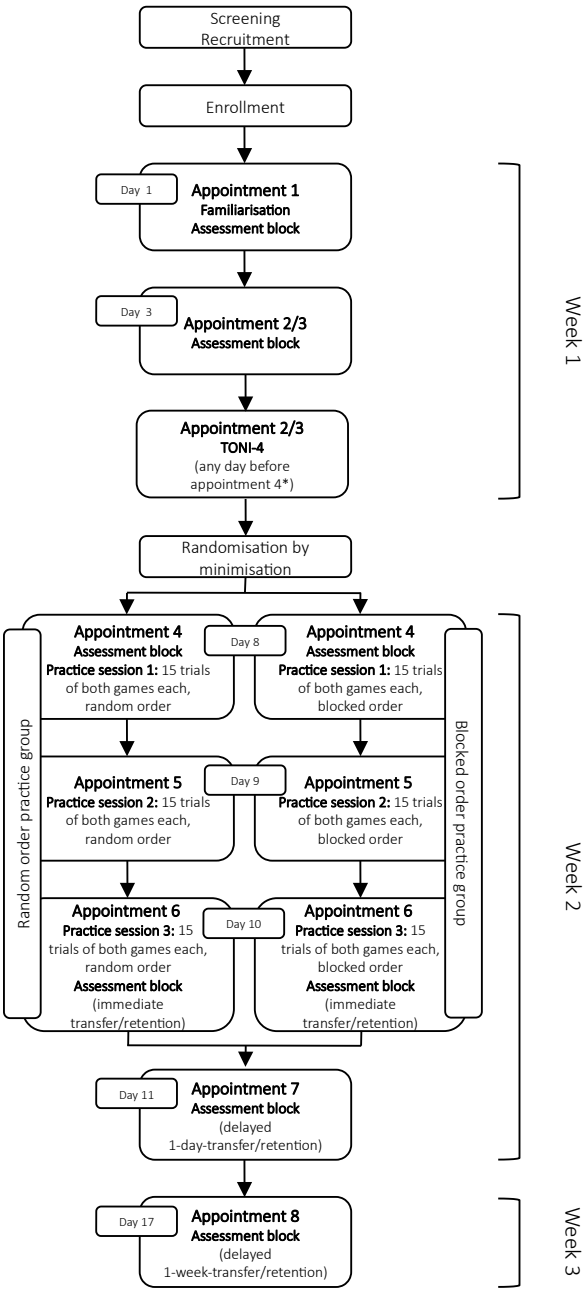


Fig. 5.5. Study procedure. The study duration for each participant is 3 weeks. Week 1 contains the familiarisation of the two versions of the exergame, two assessment blocks (Melbourne Assessment 2 and exergame tests) and the Test of Nonverbal Intelligence, Fourth edition (TONI-4). The TONI-4 will be planned before the first practice session day (i.e. on appointment 2 or 3). Week 2 contains three practice sessions with an assessment block proceeding the first practice session and another assessment block after the last practice session to evaluate immediate transfer and retention. One day after the last practice session, the assessment block is repeated to assess the delayed one-day transfer and retention). During week 3, the last assessment block is scheduled one week after the last practice session to assess the delayed 1-week transfer and retention. Abbreviations: TONI-4 = Test of Nonverbal Intelligence, Fourth edition.

Week 2

On each of three consecutive days, all participants will practice 30 trials (15 trials of the horizontal and 15 of the vertical version of the exergame) in either a blocked or random order (according to their group allocation). Right before the first and after the last practice session, an assessment block will take place to assess the acquisition and immediate transfer and retention. One day after the last practice session, the assessment block will be repeated (1 day delayed transfer and retention).

Week 3

One week after the last training session, the assessment block will be repeated (1 week delayed transfer and retention).

The first session will include robot set-up and familiarisation and will be scheduled for approximately 75 min. Assessment blocks will take approximately 35 min and practice sessions 60 to 90 min. As the duration of an exergame trial depends on the child's performance, it will vary largely and will affect the overall duration of the practice session. Once the robot settings have been adjusted to the participants' anthropometrics, no additional set up time will be required. We defined an upper time limit of 120 min per session, which could be reached during practice sessions 1 and 3 that will also include a previous or subsequent assessment block. Sessions will be aborted after 120 min.

Standardisation and blinding

All the appointments will be scheduled during the same half of the day throughout the 3 weeks to reduce the influence of daytime on the outcome (e.g. fatigue increasing throughout the day, the severity of spasticity changing over the day, etc.). A trained physiotherapist or human movement scientist will conduct the practice sessions and the assessments following standardised instructions. The same person will conduct at least all assessment blocks with a participant and, whenever possible, also the practice sessions. In each session, an assistant will note irregularities and support the therapist.

While the assessor of the primary outcome (video analysis of the $MA2_{fluency}$) will be blinded to the participants' allocation and the time-point of the assessment, the participants will be blinded to the detailed aims of the study. Both will be informed that the intervention will investigate motor learning with new exergames but not about our interest in investigating the effects of different practice orders.

Data analysis

We will present the numbers of the recruitment and measurement procedure in the CONSORT extension to randomised pilot and feasibility trials flow diagram (Figure 5.6). We will report the participant characteristics and the baseline descriptives for each participant.

ChARMin data will be processed by a custom-made Matlab algorithm (The MathWorks Inc.). Statistical analysis will be performed with IBM SPSS Statistics 24.

In case of incomplete datasets, loss to follow-up or if a participant would be treated in the other treatment group than the one he or she was allocated to, we plan an intention-to-treat analysis. In case of missing data (e.g. a participant is not able to attend a session), we plan to perform a multiple imputation by chained equation.

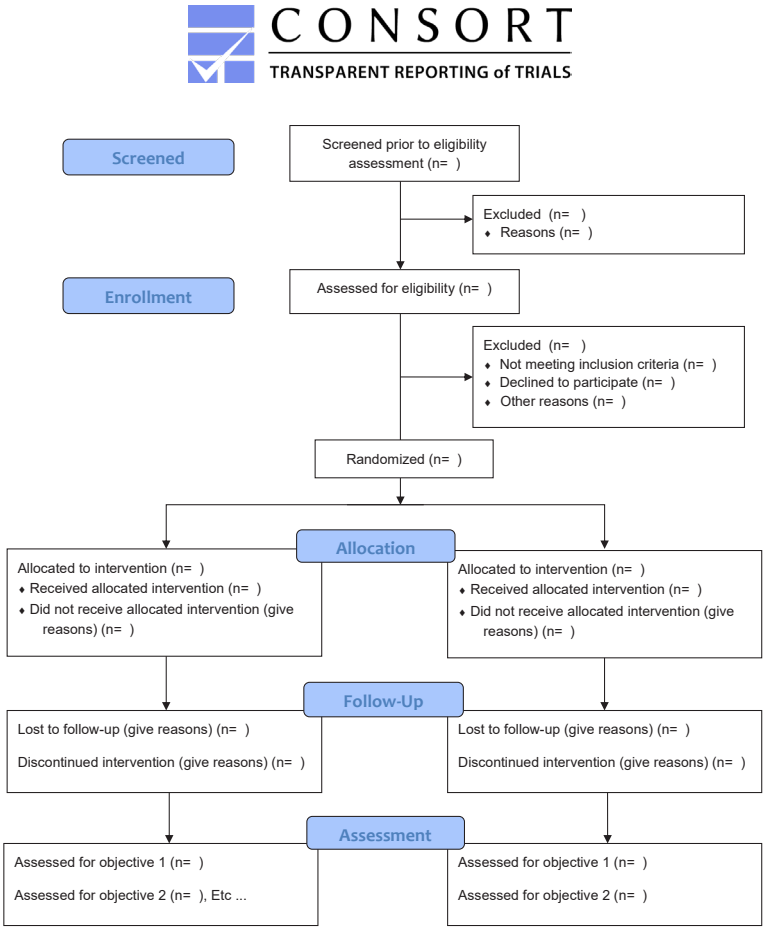


Fig. 5.6. The Consolidated Standards of Reporting Trials (CONSORT) extension to randomised pilot and feasibility trials flow diagram [27].

Analysis of feasibility

The ten areas (see outcome measures section) will be evaluated with the information obtained during the study. Only questions 8, 9 and 10 will require statistical analyses to be answered:

8. Responsiveness analysis of the outcome measures for this specific intervention: the data ($MA2_{fluency}$ and exergame test data) of the measurement time point immediately before the first and right after the last practice session will be compared. For internal responsiveness, the standardised response mean will be calculated. For the external responsiveness, a Pearson correlation coefficient will be calculated between the change-scores (i.e. between the two measurement time points) of the $MA2_{fluency}$ and the ChARMin parameter nP_{norm}.
9. Effect due to the parallel therapies within the rehabilitation setting: We will include data of all participants to test for a significant change in exergame and $MA2_{fluency}$ scores between day 1 and day 3 in week 1 (i.e. without specific intervention, see Figure 5.4) using a paired T test or Wilcoxon signed-rank test if a non-parametric analysis will be required. Assuming that the participants will have similar therapy schedules uring week 1 and 2, and no change is measured during week 1, a change during week 2 could be attributed to the robotic training. However, if outcomes will improve significantly in week 1, changes observed during week 2 will need to be interpreted with caution, as they cannot be attributed to the study's intervention alone.
10. The sample size estimation for the main trial will be calculated according to the formula

$$n = \frac{\left(Z_\alpha + Z_\beta\right)^2 \times 2\sigma^2}{(\mu_1 - \mu_2)^2}$$ [46].

Where n=sample size, Z_α=standard normal z-value for a significance level α=0.05, which is 1.96, Z_β=standard normal z-value for the power of 80%, which is 0.84. The pooled standard deviation of the pre-and post-intervention differences (both groups) is indicated by σ, μ_1 is the mean pre-post-intervention difference of the intervention group 1, μ_2 is the mean pre-post-intervention difference of the intervention group 2.

Analysis of motor learning outcomes of the main trial

As recommended by the CONSORT extension to pilot and feasibility studies, we will report 95% confidence interval estimates of the motor learning outcomes [27], as the study will be most likely underpowered. Hypothesis testing will only be conducted if the recruited sample size is large enough to warrant sufficient statistical power.

Primary outcome

Immediate transfer: The relative difference $\left(\frac{MA2_4 - MA2_3}{MA2_3}\right)$ of the $MA2_{fluency}$ sum score between the time point right before the first practice session ($MA2_3$) and the time point immediately after the last practice session ($MA2_4$) will be compared between the groups. A *T* test (or Mann-Whitney *U* test, if data are not normally distributed) will be performed.

Secondary outcomes

For all secondary outcomes, a multifactorial repeated measures analysis of variance (ANOVA) with 'practice group' and 'time point' as factors will be conducted over all the time points. A post-hoc analysis (Bonferroni correction for multiple comparisons) will be used to test for significant changes between the appropriate time points.

Delayed transfers: The $MA2_{fluency}$ sum scores will be compared between the time point right before the first practice session, 1 day after the last session (1-day-delayed transfer), and 1 week after the last session (1-week-delayed transfer).

Immediate retention, delayed retentions, acquisition: For the several retention time points and the acquisition, the parameter nP_{norm} (obtained during the exergame tests and the practice sessions, respectively) will be evaluated. Comparisons will be made between the data obtained during the exergame tests right before the first practice session with the data from immediately after the last practice session (immediate retention), 1 day after the last practice session (1-day retention), and 1 week after the last practice session (1-week retention). For the acquisition, all the data points obtained during the practice trials will be compared.

Assumptions

Generally, statistical testing is only performed in case of sufficient statistical power. For a T test, the normal distribution of the data will be required, which will be tested with a Shapiro-Wilk test. For an ANOVA, the homoscedasticity of values of the groups and normal distribution of the population of the residues will be tested with the Levene test and the Shapiro-Wilk test, respectively.

DISCUSSION

We aimed to present the study protocol of a randomised controlled single-blinded pilot trial to explore the feasibility of testing the contextual interference effect while practicing robot-assisted upper limb tasks with exergames in a representative sample of children with congenital or acquired brain lesions.

Putting the pilot study in the wider context

It is common that different motor tasks are practiced within the same therapy session in paediatric neurorehabilitation. Human movement is variable (e.g. [19]) and variability generally seems to be beneficial for learning of new motor skills [47]. Despite this, high-level evidence about the contextual interference effect in children in general and particularly in children with brain lesions is missing [20]. If children with congenital or acquired brain lesions would respond similarly as healthy adults to the contextual interference effect, random practice would be more beneficial compared to blocked practice at retention and transfer. On the long-term, this knowledge could contribute to the optimisation of treatment protocols leading to improved clinical outcomes in children undergoing neurorehabilitation.

All rehabilitation goals are individual, yet participation [48] and independence in daily life are often mentioned as the most important aims for the rehabilitation of children [49]. The importance of generalisability of motor improvements (i.e. the result of motor learning) and the retention thereof in the daily life of the child is probably not a matter of dispute. Yet, there is a gap between the somewhat constrained reaching movement during an exergame with a robot that is limited to specific movement directions and a goal-directed reaching movement in daily life. Or, in the language of the International Classification of Functioning, Disability and Health [50], practicing 'body functions' or 'capacity' in a robot might not directly influence 'performance' in a daily life environment. While in the robotic condition, the starting position is relatively fixed and the number of movement directions is prescribed and somewhat constrained, reaching in daily life has not such limitations, which leads to many more movement opportunities [51], including numerous movement directions requiring movement control over multiple joints. Furthermore, the use of upper extremities in daily life is mostly a bimanual matter in persons without neurological deficits [52, 53]. While there will much more questions coming up along the way, a randomised controlled trial, as we described it in this protocol, might answer the first question. As there is not enough knowledge available to start with such a study, the conduct of a pilot study is inevitable. Concludingly, this protocol of the planned pilot study is the first step of a series of successive questions and, hopefully, answers, which ultimately leads to an optimised transfer of motor learning improvement into everyday life performance.

Potential risks and benefits for the participants

It has been stated that scientists need to develop sensitivity for the risks and benefits of the participants when volunteering for a trial [54]. Currently, we do not know whether the blocked or random practice order would be more beneficial. Nevertheless, we assume that the practice schedule described in this study protocol could lead to an improvement in motor functions. Despite that participants will practice only three sessions, according to literature, the resulting 1440 repetitions

should induce motor learning as 500 repetitions led to motor learning effects in adults after stroke [37]. Therefore, for all the children participating in this study, a direct benefit can be expected. Especially for such a vulnerable population as children with neurological impairments, the risks of participating in a study should be minimal or, if greater, should be outweighed by the anticipated benefit [55]. Our study provides a safe setting, and we expect no severe adverse events. Potential risks might be fatigue, muscle soreness, a more challenging rehabilitation schedule due to a higher number of sessions than usual and skin irritations due to pressure and friction of the cuffs. The risk-benefit trade-off of this pilot study seems balanced and therewith justifiable.

CONCLUSION

After the execution of this pilot study, we will be able to decide upon the questions whether the main randomised controlled trial could be feasible and, if yes, how it should be performed.

ACKNOWLEDGEMENTS

We would like to thank Anja Gut, research associate, for her assistance, Andrina Kläy, research associate, for her technical support, Jeffrey Keller, PhD student, for the statistical advice, and Susanne Staubli for the graphic realisation.

AUTHORS' CONTRIBUTIONS

JVG conceptualised and designed the study protocol and drafted the manuscript. CHGB supported the conceptualisation of the study protocol and substantively revised the manuscript. UK created the new software that is described in this study protocol. HJAvH supported the conceptualisation of the study protocol, substantively revised the manuscript, and obtained the funding. All authors have approved the submitted version and have agreed both to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which they were not personally involved, are appropriately investigated, resolved and the resolution documented in the literature.

FUNDING

This study protocol has been supported by the Clinical Research Priority Programme, Neuro-Rehabilitation of the University of Zurich, and the Mäxi Foundation.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study has been approved as part of the ChARMin project (sub-project 5: Motor learning) by the Ethics Committee of the Canton Zurich (BASEC-Nr. PB_2016-02450) and the Swiss Agency for Therapeutic Products (Swissmedic reference number: 2015-MD-0009). According to the Ethics Committee's guidelines, all the participants should give their verbal consent; children of 15 years and older and all the legal representatives should give informed written consent. All the participants and their legal representatives have the right to withdraw their participation at any time.

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

1. Sellier E, PLatt MJ, Andersen GL, Krägeloh-Mann I, De la Cruz J, Cans C, et al. Decreasing prevalence in cerebral palsy: a multi-site European populationbased study, 1980 to 2003. Dev Med Child Neurol. 2015;58:85–92.

2. Olsen M, Vik A, Lund Nilsen TI, Uleberg O, Moen KG, Fredriksli O, et al. Incidence and mortality of moderate and severe traumatic brain injury in children: a ten year population-based cohort study in Norway. Eur J Paediatr Neurol. 2019;23:500–3. Mallick AA, O'Callaghan FJK. The epidemiology of childhood stroke. Eur J Paediatr Neurol. 2010;14:197–205.

4. Ball WS, Huisman TAGM. Congenital and acquired brain abnormalities of the pediatric brain. In: Hodler J, von Schulthess GK, Zollikofer CL, editors. Diseases of the brain, head and neck, spine 2012-2015. Milano: Springer; 2012. p. 188–202.

5. Zwicker JG, Harris SR. A reflection on motor learning theory in pediatric occupational therapy practice. Can J Occup Ther. 2009;76(1):29–37.

6. Kitago T, Krakauer JW. Motor learning principles for neurorehabilitation. In: Barnes MP, Good DC, editors. Handbook of Clinical Neurology. Amsterdam: Elsevier; 2013. p. 93–103.

7. Krakauer JW. Motor learning: its relevance to stroke recovery and neurorehabilitation. Curr Opin Neurol. 2006;19(1):84–90.

8. Schmidt RA, Lee TD. Motor control and learning: a behavioral emphasis. 4th ed. Human Kinetics: Champaign; 2005.

9. Maas E, Robin DA, Hula SNA, Freedman SE, Wulf G, Ballard KJ, et al. Principles of motor learning in treatment of motor speech disorders. Am J Speech-Language Pathol. 2008;17(3):277–98.

10. Battig W. Intratask interference as a source of facilitation on transfer and retention. In: Voss J, editor. Topics in learning and performance. New York: Academic Press; 1972. p. 131–59.

11. Magill RA, Hall KG. A review of the contextual interference effect in motor skill acquisition. Hum Mov Sci. 1990;9(3–5):241–89.

12. Hodges NJ, Lohse KR, Wilson A, Lim SB, Mulligan D. Exploring the dynamic nature of contextual interference: previous experience affects current practice but not learning. J Mot Behav. 2014;46(6):455–67.

13. Shea JB, Zimny S. Context effects in memory and learning movement information. In: Magill, Richard A, editor. Memory and control of action. Amsterdam: North-Holland; 1983. p. 345–66.

14. Lage GM, Ugrinowitsch H, Apolinário-Souza T, Mário Vieira M, Albuquerque

MR, Novellino BR. Repetition and variation in motor practice: a review of neural correlates. Neurosci Behav Rev. 2015;57:132–41.

15. Lee TD, Magill RA. The locus of contextual interference in motor-skill acquisition. J Exp Psychol Learn Mem Cogn. 1983;9(4):730–46.
16. Battig W. The flexibility of human memory. In: Cermak L, Craik F, editors. Levels of processing in human memory. Hillsdale; 1979. p. 23–44.
17. Schmidt RA. A schema theory of discrete motor skill learning. Psychol Rev. 1975;82:225–60.
18. Schmidt RA. Motor control and learning: a behavioral emphasis. 5th ed. Champaign: Human Kinetics; 1988.
19. Bernstein NA. The co-ordination and regulation of movements. Oxford: Pergamon Press; 1967.
20. Graser JV, Bastiaenen CH, van Hedel HJA. The role of the practice order: a systematic review about contextual interference in children. PLoS One. 2019;14(1).
21. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: a systematic review. Lancet Neurol. 2009;8(8):741–54.
22. van Hedel HJ, Aurich-Schuler T. Clinical application of rehabilitation technologies in children undergoing neurorehabilitation. In: Reinkensmeyer DJ, Dietz V, editors. Neurorehabilitation technology. 2nd editio. Switzerland: Springer International Publishing; 2016. p. 283–310.
23. Oh Y, Yang S. Defining exergames and exergaming. In: Meaningful Play 2010. 2010.
24. Bonney E, Jelsma D, Ferguson G, Smits-Engelsman B. Variable training does not lead to better motor learning compared to repetitive training in children with and without DCD when exposed to active video games. Res Dev Disabil. 2017;62:124–36.
25. Prado MTA, Gonçalves Luiz Fernani DC, Dias da Silva T, Smorenburg ARP, de Abreu LC, Bandeira de Mello Monteiro C. Motor learning paradigm and contextual interference in manual computer tasks in indivisuals with cerebral palsy. Res Dev Disabil. 2017;64:56–63.
26. Thabane L, Lancaster G. A guide to the reporting of protocols of pilot and feasibility trials. BMC Pilot Feasibility Stud. Pilot and Feasibility Studies. 2019; 8:5–7.
27. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials.

BMJ. 2016;355:i5239.

28. Chan A-W, Tetzlaff JM, Altman DG. SPIRIT 2013 Statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7.

29. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Ohrvall A-M, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: Scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006;48(7):549–54.

30. Altman DG, Bland JM. Treatment allocation by minimisation. Br Med J 2005; 330(April):843.

31. Ritter N, Kilinc E, Navruz B, Bae Y. Test review (TONI-4). J Psychoeduc Assess. 2011;29(5):484–8.

32. Keller U, Van Hedel HJA, Klamroth-Marganska V, Riener R. ChARMin: the first actuated exoskeleton robot for pediatric arm rehabilitation. IEEE/ASME Trans Mechatronics. 2016;21(5):2201–13.

33. Keller U, Schölch S, Albisser U, Rudhe C, Curt A, Riener R, et al. Robot-assisted arm assessments in spinal cord injured patients: a consideration of concept study. PLoS One. 2015;10(5).

34. Wood CA, Ging CA. The role of interference and task similarity on the acquisition, retention, and transfer of simple motor skills. Res Q Exerc Sport. 1991;62(1):18–26.

35. Granda Vera J, Montilla MM. Practice schedule and acquisition, retention, and transfer of a throwing task in 6-yr.-old children. Percept Mot Skills. 2003; 96(3 Pt 1):1015–24.

36. Lang CE, MacDonald JR, Reisman DS, Boyd L, Jacobson Kimberley T, Schindler-Ivens SM, et al. Observation of amounts of movement practice provided during stroke rehabilitation. Arch Phys Med Rehabil. 2009;90(10): 1692–8.

37. Boyd L, Winstein CJ. Explicit information interferes with implicit motor learning of both continuous and discrete movement tasks after stroke. J Neurol Phys Ther. 2006;30(2):46–57.

38. Keller JW, van Hedel HJ. Weight-supported training of the upper extremity in children with cerebral palsy: a motor learning study. J Neuroeng Rehabil. 2017;14:87.

39. Randall M, Johnson L, Reddihough D. The Melbourne assessment 2 [Internet]. [cited 2019 Mar 29]. Available from: https://www.rch.org.au/melbourneassessment/.

40. Wang T-N, Liang K-J, Liu Y-C, Shieh J-Y, Chen H-L. Psychometric and clinimetric

properties of the Melbourne Assessment 2 in children with cerebral palsy. Arch Phys Med Rehabil. 2017;98(9):1836–41.

41. The Royal Children's Hospital Melbourne. The Melbourne assessment 2—scoring criteria.

42. Kandel S, Perret C. How do movements to produce letters become automatic during writing acquisition? Investigating the development of motor anticipation. Int J Behav Dev. 2015;39(3):113–20.

43. Meulenbroek RGJ, van Galen GP. The production of connecting strokes in cursive writing: Developing co-articulation in 8 to 12 year-old children. In: Plamondon R, Suen CY, Simner ML, editors. Computer recognition and human production of handwriting. Singapore: World Scientific; 1989. p. 105–18.

44. Hoffmann TC, Glasziou PP, Milne R, Perera R, Moher D, Altman DG, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. Br Med J. 2014;1687.

45. Randall M, Imms C, Carey L, Pallant J. Rasch analysis of The Melbourne Assessment of unilateral upper limb function. Dev Med Child Neurol. 2014; 56:665–72.

46. nQuery Powering Sample Size [Internet]. [cited 2019 Apr 16]. Available from: https://www.statsols.com/how-to-use-a-sample-size-calculator.

47. Dhawale AK, Smith MA, Ölveczky BP. The role of variability in motor learning. Annu Rev Neurosci. 2017;40:479–98.

48. McCarron RH, Watson S, Gracey F. What do kids with acquired brain injury want? Mapping neuropsychological rehabilitation goals to the International Classification of Functioning, Disability and Health. J Int Neuropsychol Soc. 2019;25(4):403–12.

49. Angeli JM, Harpster KL, Hanson E, Sheehan A, Schwab SM. Patient- and caregiver-identified preferences: dimensions of change in developmental therapy treatment goals. Dev Neurorehabil. 2019;22(1):39–46.

50. World Health Organization. International classification of functioning, disability and health—children & youth version. Geneva, CH: WHO Press, World Health Organization; 2007.

51. Rösblad B. Reaching and eye-hand coordination. In: Henderson A, Pehoski C, editors. Hand function in the child: foundations for remediation. 2nd editio. St. Louis: Mosby Elsevier;2006. p. 89-100.

52. Kilbreath SL, Heard RC. Frequency of hand use in healthy older persons. Aust J Physiother Elsevier. 2005;51(2):119–22.

53. Bailey RR, Klaesner JW, Lang CE. Quantifying real-world upper limb activity in nondisabled adults and adults with chronic stroke. Neurorehabil Neural Repair. 2015;29(10):969–78.

54. ALLEA ALL Eurpean Academies, Hermerén G, Benaki A, Fagot-Largeault A, Trontelj J, Strøm Bull K, et al. Ethics education in science—statement by the ALLEA permanent working group on science and ethics. Germany; 2013 p. 1–16.

55. Nelson RM. Ethik in klinischen Studien mit Kindern- die amerikanische Sichtweise. In: Brochhausen C, Seyberth HW, editors. Kinder in klinischen Studien - Grenzen medizinischer Machbarkeit? Münster: LIT Verlag; 2005. p. 181–92.

CHAPTER 6

CONTEXTUAL INTERFERENCE IN CHILDREN WITH BRAIN LESIONS: A PILOT STUDY INVESTIGATING BLOCKED VS. RANDOM PRACTICE ORDER OF AN UPPER LIMB ROBOTIC EXERGAME

Judith V Graser
Caroline HG Bastiaenen
Anja Gut
Urs Keller
Hubertus JA van Hedel

Under review

ABSTRACT

Introduction:

Evidence about contextual interference in children with brain lesions when practicing motor tasks is lacking. Our main objective was to evaluate the feasibility of a randomised controlled trial (RCT) comparing blocked with random practice order of an upper limb robotic exergame to improve reaching in children with neuromotor disorders with a pilot trial.

Methods: We recruited children with brain lesions and impaired upper limb functions who underwent a three-week schedule that consisted of baseline assessments, intervention period (participants were randomised to a blocked or random order group), and follow-up assessment. We evaluated ten feasibility criteria, including the practicability of the inclusion/exclusion criteria, recruitment rate, feasibility of randomisation, scheduling procedure, and the participants' programme adherence.

Results: The inclusion/exclusion criteria were not completely feasible as patients who were not able to perform the exergames were included. Twelve participants were recruited, six datasets were used for analysis. The scheduling and randomisation procedures were generally feasible, but the procedure was only partially feasible for the participants, as some sessions were aborted due to lack of motivation and fatigue.

Conclusion: An RCT following this study protocol is not feasible. We formulated suggestions for future studies that aim to investigate contextual interference as in this pilot study.

Trial registration: ClinicalTrials.gov Identifier: NCT02443857, registered May, 14th 2015
https://clinicaltrials.gov/ct2/
results?cond=&term=NCT02443857&cntry=&state=&city=&dist=

INTRODUCTION

Motor learning forms the basis of neurorehabilitation [1]. The initial phase during which motor skills are gained is referred to as acquisition. Nevertheless, it is important for successful neurorehabilitation that a practiced motor skill can be maintained over time and generalised to another setting or situation. Hence, the aspects of motor learning that are of particular interest for neurorehabilitation are retention and transfer. We need a better understanding of motor learning to integrate underlying principles such as different types of practice or feedback successfully in, for example, hand rehabilitation [2].

Motor learning requires practice, consisting of a high number of repetitions but not of the exact same movement [3], as practice should include some variability ("repetition without repetition" [4]). It has been proposed that motor learning is generated when motor problems need to be solved [2]. Considering this, "repetition without repetition" is not repeating the solution for a specific motor problem but repeating the process of solving the problem by improving it with every trial [4]. Performing a large number of repetitions can be tedious, especially for children. Hence, another reason for implementing variable practice is to influence the motivation needed to perform so many repetitions constantly in a positive way.

Variable practice is understood as practicing different variations of the task. An effect occurring when practicing different tasks within one therapeutic session is contextual interference [5]. The contextual interference effect in motor learning refers to the interference that results from practising a task within the context of other tasks in a practice session. A continuum of increasing variability lead to different levels of contextual interference. When variability is low, for example, when tasks are practiced in blocked order, i.e. several repetitions of one task before switching to the next task (e.g. AAA…, BBB…, CCC…), the level of contextual interference is low [6]. When variability is high, for example, if tasks are performed in random order (e.g. ABC…, CCA…, BAB…), participants practise under conditions of high contextual interference [6]. In healthy adults, practicing with low contextual interference leads to improved acquisition but reduced transfer and retention compared to practicing with high contextual interference [6].

Three different hypotheses underlying this effect have been discussed: the elaboration hypothesis, the forgetting-reconstruction hypothesis, and the retroactive inhibition hypothesis. The elaboration hypothesis suggests that practicing skills in random order results in more distinctive processing and is beneficial for transfer and retention [7,8]. The forgetting-reconstruction hypothesis is based on the idea that practicing in random order requires the learner to forget after each task what he or she learned in order to focus on the next task [8]. This might disturb performance during acquisition but is beneficial for retention [8]. The retroactive inhibition hypothesis [9–11] suggests that the contextual interference effect occurs due to a

disadvantage of blocked practice and not due to an advantage of random practice. Blocked practice order might inhibit the previous memory leading to differences in performance between the random and blocked practice [10].

In children and youths practicing typical sports skills in a field setting under high contextual interference (random) or mixed (random and blocked) conditions, a systematic review and meta-analysis reported effect sizes between -0.8 and 0.3 for transfer or retention measures [12]. However, the difference between blocked and random practice order was not evaluated. Another systematic review included 25 papers investigating contextual interference in typically developing children and children with brain lesions [13]. No conclusion about the contextual interference effect could be made as results were diverging, the methodological quality of these studies was low, and the risk of bias high [13].

In recent years, rehabilitation technologies, such as robot-assisted training devices, have been applied increasingly, particularly in paediatric neurorehabilitation [14,15]. Robot-assisted devices in combination with exergames provide playful exercise, which increases engagement and motivation to achieve larger numbers of movement repetitions [16]. Most rehabilitation technologies contain sensors that can be used to estimate the status and functional improvement of patients objectively [17]. Applying such technologies to investigate contextual interference in children with brain lesions seemed therefore a promising option. One recent study applied a computer maze game and compared constant practice order (total of 30 trials of one maze) with random practice order (total of 30 trials of five different mazes) in children with cerebral palsy (CP) and typically developing (TD) children [18]. The random order group performed the maze games faster at retention (TD and CP combined) and transfer (CP and TD combined but also the single groups) [18]. Usually, contextual interference studies include blocked and random order groups that practice the same number of repetitions of each task. This study used constant and random practice order, and group differences might be related to the number of different mazes practiced in the two groups (one vs. five) and not the practice order per se. Hence, scientific experience in this specific field remains scarce.

It might be difficult to conduct such a study within a clinical setting for various reasons, including restricted numbers of eligible participants. To avoid unnecessary waste of resources, it is recommended to perform a pilot study preceding the final randomised clinical trial (RCT) [19]. Therefore, our main objective was to evaluate the feasibility of a randomised controlled intervention trial assessing the contextual interference effect by comparing blocked with random order practice of an exergame played with a robotic upper limb exoskeleton device to improve reaching movements in children with neuromotor disorders.

METHODS

We designed this randomised controlled pilot study with two intervention arms in line with the Consolidated Standards of Reporting Trials (CONSORT) statement extension for randomised pilot and feasibility trials [20], the recommendations about conducting a pilot study provided by Thabane et al., 2010 and the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) [21]. Methodological details were described in the study protocol [22].

Participants

We included children meeting the following inclusion criteria:

- Age between five and 18 years
- Congenital or acquired (subacute or chronic state, i.e. at least three months after the incident) brain lesion
- Uni- or bilaterally impaired upper limb function (spasticity, dyskinesia or mixed conditions) Undergoing inpatient neurorehabilitation
- Manual Ability Classification System (MACS) level I (objects are handled easily and successfully) to IV (inability to handle objects or to complete even simple actions with the hands) [23]
- Ability to sit for approximately 60 minutes without lateral trunk support
- Ability to understand and follow test instructions
- Ability to visually follow events on a screen in approximately one meter distance
- Ability to communicate discomfort or pain

Children were excluded if meeting the following exclusion criteria:

- Surgery or upper limb Botox injections during the past six months
- Upper limb skin lesions
- Severe obesity
- Severe spasticity (Modified Ashworth Scale >3, i.e. considerable increase in muscle tone, passive movement difficult, rigidity [24])
- Fixed contractures or deformities of the upper limbs.

Recruitment

Each week, a therapist screened the patients entering the Swiss Children's Rehab of the University Children's Hospital Zurich in Affoltern am Albis, Switzerland. Potential participants were informed about the study. Written informed consent was obtained

from all the participants' legal representatives and participants aged 15 years and older, who agreed to participate. All participants provided verbal consent. This study was approved by the Ethics Committee of the Canton of Zurich (BASEC-Nr. PB_2016-02450, subproject 5 motor learning) and the Swiss Agency for Therapeutic Products (Swissmedic reference number: 2015-MD-0009). We enrolled participants during approximately one year, between the end of June 2018 and June 2019.

Group allocation

Participants were allocated to one of the two intervention groups (blocked or random practice order) by randomisation through minimisation, which is a method ensuring the balance of several prognostic factors between study groups in small samples [25]. Minimisation parameters were the following descriptors: age (preschool age: 5-6 years, primary school age: 7-12 years, and secondary school age and older: 13-18 years); gender (female, male); diagnosis (congenital, acquired), manual ability (MACS level I, II, III, or IV), and cognitive ability according to the Test of Nonverbal Intelligence – Fourth edition (TONI-4), which was applied by a trained neuropsychologist and evaluates abstract reasoning and problem solving (Index Score <70: very poor, 70-79: poor, 80-89: below average, 90-110: average, 111-120: above average, 121-130: superior, and >130: very superior) [26].

When the principal investigator obtained these data, she informed the study nurse that the data were ready for the minimisation process. Minimisation is a secure method when conducted by an independent person [25]. Therefore, the study nurse person who performed the randomisation and allocation did not participate in any other study procedure and was unaware of the aim of the study. She added the information in the custom-made Matlab script. In case of absence, a back-up person was assigned, who was not involved in other study operations. Participants could not influence the choice in which group they would practice.

Robot

For this pilot study, we chose the ChARMin robot [27]. In short, ChARMin is an exoskeleton robot for the rehabilitation of upper limb function in children from five to 18 years. It can be adjusted to the anthropometric sizes of younger and older children as two different distal modules were developed (Figure 6.1). ChARMin can actively support arm movements while playing various audio-visual exergames shown on the screen in front of the participant.

Exergames

Contextual interference is assumed to occur when different motor programs are involved in different motor tasks [5]. It has been proposed that spatial and topographical characteristics are invariant within the same motor programs [4]. Thus, choosing different spatial characteristics (i.e. moving in different planes) ensures the involvement of different motor programs. Therefore, we designed a

specific ChARMin exergame, which can be played in the frontal or the transversal plane (Figure 6.1).
In the exergame, participants had to control an avatar on the screen by changing the position of the hand while being attached to the exoskeleton. Eight different targets, which appear in random order radially around a centre position, had to be reached. The children were instructed to reach the targets 'using the most direct path and as fast as possible'. The exergame was played in the frontal plane (30 cm or 35 cm in front of the shoulder joint with the small or the large distal module, respectively) and in the transversal plane (10 cm below the shoulder joint with both distal modules) (Figure 6.1). Participants performed the reaching movements within and restricted to these planes. Each exergame version provided two scenarios (transversal plane: "Unicorn", "Snail"; frontal plane: "UFO", "Submarine") which were shown on the screen (Figure 6.2).

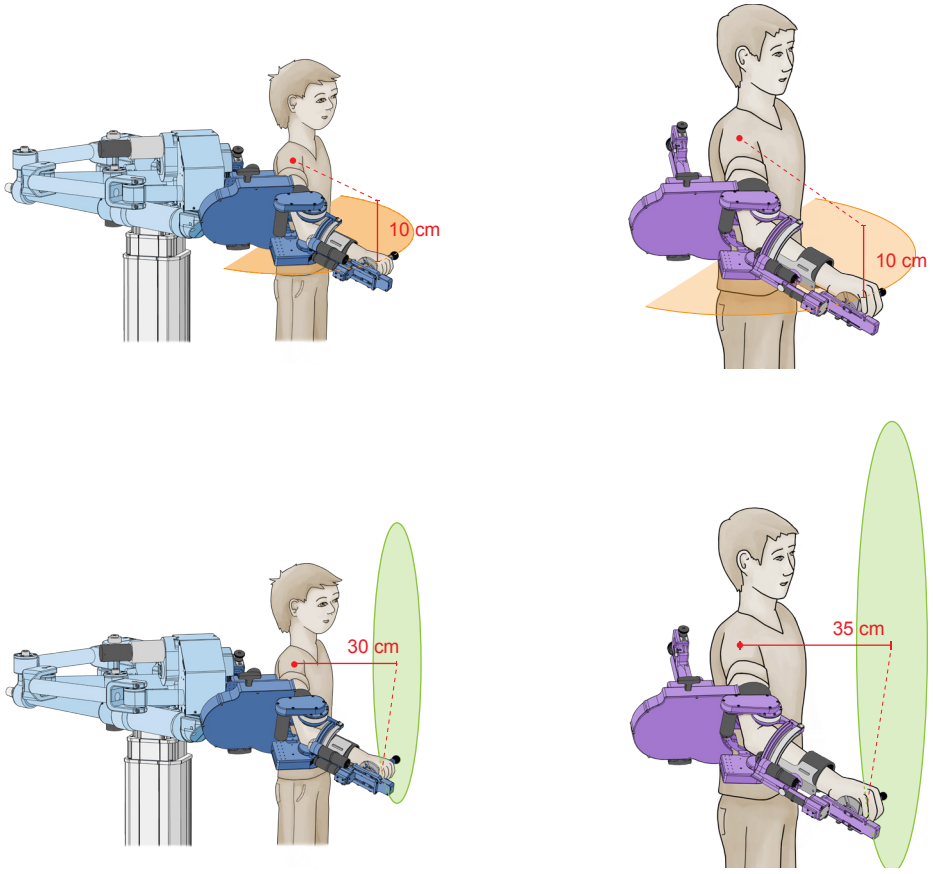


Fig. 6.1. ChARMin with two distal modules and different movement planes. A1 Small distal module (dark blue), movements restricted to the transversal plane. A2 Large distal module (purple), movement direction restricted to the transversal plane. B1 Small distal module (dark blue), movements restricted to the frontal plane. B2 Large distal module (purple), movements restricted to the frontal plane.

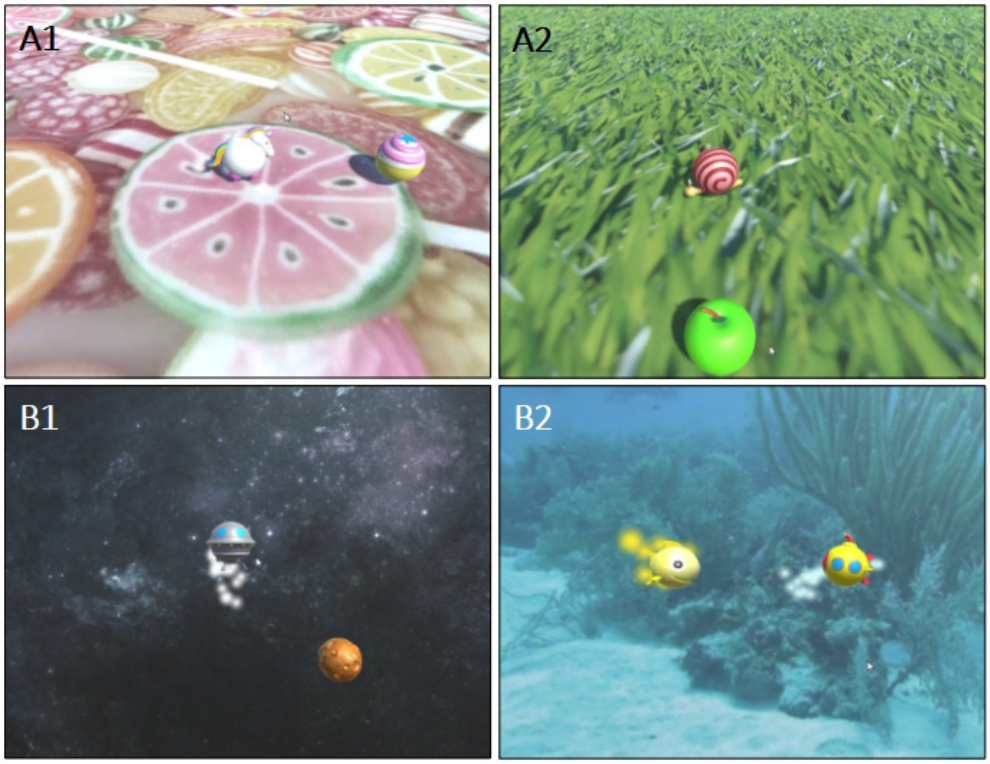


Fig. 6.2. Exergame scenarios displayed on the computer screen. Each exergame version provided two different scenarios. A1 exergame version transversal plane, scenario "Unicorn"; A2 exergame version transversal plane, scenario "Snail"; B1 exergame version frontal plane, scenario "UFO"; B2 exergame version frontal plane, scenario "Submarine".

Interventions

Participants of the blocked order group performed all the trials of one exergame version (e.g. in the frontal plane) before switching to the second exergame version (e.g. transversal plane). The participants in the random order group performed the trials of both exergame versions in a pseudo-random order, as we restricted the number of consecutive repetitions of one exergame version to two to keep the contrast between the interventions large. The participants were free to choose between the scenarios for each trial.

The participants trained with the more affected upper limb. If both arms were affected equally, the arm subjectively used as dominant arm in daily life was chosen for practice.

Study procedure

Each participant followed a three weeks schedule. Week 1 was designed as a control week to evaluate potential changes occurring without the study-specific training and compare them with the changes occurring during the practice week.

Week 1 contained two sessions of assessments (exergame test and the Melbourne Assessment 2, subscale fluency ($MA2_{fluency}$); for descriptions of the assessments, see below), two days apart, without actually practising with the robot. During week 1, as soon as all the descriptors were obtained, group allocation by randomisation through minimisation was performed. Week 2 was the practice week where three practice sessions took place on three consecutive days. The participants practiced both exergame versions in either blocked or random order, depending on their group allocation. The number of repetitions was set to 30 per exergame version. Before the first and immediately after the last session, an assessment block consisting of the exergame test and the $MA2_{fluency}$ was performed. The same assessment block was repeated one day after the last practice session. Week 3 was the follow-up week with an assessment block one week after the last practice session. All sessions (practice and assessments) of the same participant were planned on the same half of day. Figure 6.3 shows the SPIRIT Figure, which contains all the study appointments chronologically ordered.

	Study period								
	Enrolment	**Allocation**	**Post-allocation**						**Close-out**
TIMEPOINT	$-t_1$		t_1	t_3	t_8	t_9	t_{10}	t_{11}	t_{18}
ENROLMENT:									
Eligibility screen	X								
Informed consent	X								
Recording of minimisation parameters	X								
Allocation		X							
INTERVENTIONS:									
Familiarisation			X						
Blocked order practice					X	X	X		
Random order practice					X	X	X		
ASSESSMENTS:									
Test of Non-verbal Intelligence	X								
Melbournbe Assessment 2 (subscale «fluency»)			X	X	X		X	X	X
Exergame test			X	X	X		X	X	X

Fig. 6.3. The Standard Protocol Items for Clinical Trials (SPIRIT)-figure: schedule of enrolment, interventions, and assessments. Abbreviations: t = time point.

Outcome measures

We differentiated between three phases of learning: transfer, retention, and acquisition.

The transfer was measured with the $MA2_{fluency}$, which assesses the movement quality in children [28] and is valid, reliable, and responsive in children with cerebral palsy [29]. The sum score of the $MA2_{fluency}$ ranges between zero and 21 points as the seven items are rated with zero to three points. Higher numbers indicate better test performance. An experienced physiotherapist or health scientist administered the MA2 according to the test manual. The $MA2_{fluency}$ was videotaped. An experienced occupational therapist rated the videos of each participant after the child had completed the entire study procedure. The therapist was unaware of the time point when the videos were recorded, as the date and time stamps were deleted.

The retention was measured with data obtained during the exergame tests, which consisted of three trials of each version of the exergame repeated in blocked order. The scenario chosen initially for each exergame version was kept the same throughout all the exergame tests. We chose the parameter 'number of speed peaks normalised to the actual covered distance' (nP_{norm}). The parameter nP_{norm} is calculated as the number of velocity peaks (i.e. when the difference between a local speed minimum and a local speed maximum exceeds a value of 2.5% of the maximal speed) normalised to the covered distance of the movement path. The parameter nP_{norm} was calculated for each movement (i.e. one value from the centre to the target and one for the return to the centre starting point), leading to 16 data points per trial. As nP_{norm} quantifies movement fluency, it is supposed to reflect a similar construct as the $MA2_{fluency}$. Low nP_{norm} values refer to more fluent movement performance, and nP_{norm} shows very high relative and acceptable absolute reliability [30].

The acquisition was assessed as the change over time during all three practice sessions. To quantify the acquisition, we chose the parameter nP_{norm} calculated for each movement over the exergame trials of the three practice sessions.

Primary outcome

For the primary outcome of the trial, we chose the immediate transfer of the practiced skill with the sum score of the $MA2_{fluency}$ (i.e. the relative change between the time points immediately before and immediately after the practice session).

Secondary outcomes

We chose seven secondary motor learning outcomes. The one-day delayed transfer and the one-week-delayed transfer were measured with the $MA2_{fluency}$. The immediate retention, 1-day-delayed retention, 1-week-delayed retention, and the acquisition were measured with the nP_{norm}.

Initially, we had planned to investigate the contextual interference effect by performing various statistical analyses to determine immediate and delayed transfer, retention, and acquisition (for the detailed study protocol see [22]). However, it is not recommended to focus in a pilot study on statistical analyses if the power is insufficient [19], therefore, we presented observations of motor learning in an exploratory manner (Additional file 6.1).

Feasibility criteria

We formulated ten criteria a-priori and selected appropriate methods to verify the criteria to investigate whether the study protocol would be feasible. We determined the criteria based on Thabane et al. [19], who suggested four categories covering various reasons for conducting a pilot study: process, resources, management, and scientific. We chose and adapted the criteria according to our setting. While we ordered them chronologically in the study protocol [22], here we chose to list them according to their importance, starting with the most important.

1. Are the in- and exclusion criteria specific enough to recruit participants, suitable for this study? The main investigator noted the number of participants who were eligible and recruited but not able to perform any of the study appointments due to limitations that were not covered by the in- and exclusion criteria.
2. Is the recruitment rate feasible? We defined a fixed recruitment period and calculated the ratio between recruited participants and complete datasets that were obtained within this period.
3. Is it feasible to conduct a future trial with respect to the sample size calculated from the data obtained for the primary outcome? Data for calculating a sample size for a future randomised controlled trial could not be provided by the literature as the $MA2_{fluency}$ has not been used solely as an outcome measure before in a comparable study population. Therefore, obtaining data to perform a reliable sample size calculation was another important issue we investigated in this pilot study. We calculated the sample size based on the data obtained from the primary outcome.
4. Is the whole procedure feasible for the participants? Participants were regularly asked whether they needed a break or whether they were able to go on with the practice sessions. Based on notes from the lab journal, we evaluated the number of aborted sessions and complete terminations due to overload of the participants.
5. Is the scheduling procedure feasible? The investigator noted the number of appointments, which were not plannable or re-plannable after an unexpected change of schedule.
6. Is the randomisation procedure feasible? The number of times in which the

randomisation process was not finished on time (i.e. before the first practice sessions) was recorded. We also recorded any issues that occurred during the randomisation procedure (e.g. program error, unclear where to find the according lists, etc.).

7. Is the handling of the robot feasible? Based on notes from the lab journal, we evaluated the technical issues that had led to aborted sessions and trials.
8. Is the processing of the large amount of data feasible? The occupational therapists evaluating the primary outcome measure ($MA2_{fluency}$) was asked to report the time generally needed to evaluate six $MA2_{fluency}$ videos of one participant. Furthermore, each exergame trial generated one data file which needed to be processed with a Matlab script. The appropriate parameters (16 per trial) needed to be exported from the processed data and added in a general data table. The time it took for processing one data file until it was placed in the main data table was recorded for one example data and interpolated for the whole amount of data.
9. Are the outcome measures responsive within this setting? The data obtained during the $MA2_{fluency}$ and exergame tests at the time points immediately before and after the practice sessions were used to evaluate internal and external responsiveness. Internal responsiveness measures the ability of an assessment to change over a predefined time period while external responsiveness characterises the extent to which a change measured with the assessment of interest corresponds with the change measured by another tool [31]. We chose this approach to have both, a responsiveness outcome that reflects more the specific context (internal responsiveness) and one that has meaning in a broader range of settings (external responsiveness) [31].
10. Do parallel therapies within the rehabilitation setting influence the results? Assuming that during the first week (normal rehabilitation programme, no additional exergame practice), the change measured with the assessments would be lower than during the practice week, we defined the first week as control week. The data obtained during the $MA2_{fluency}$ and exergame tests at the two time points during the first week were used to evaluate the change within the control week in the whole sample. The participants will have similar therapy schedules during week 1 and 2, hence, if no change is measured during week 1, a change during week 2 could at least be partly attributed to the robotic training.

Target sample size

As the recruitment rate was one of the feasibility criteria, and we also aimed to calculate a sample size for a main study with data obtained from this pilot study, we recruited and enrolled eligible children for one year. Based on the number of children undergoing rehabilitation in our centre during the past years (approximately

200 per year) and considering the inclusion and exclusion criteria, we estimated to include 20 children (10 per group) in this pilot study.

Statistical analyses

Statistical analyses were performed with IBM SPSS Statistics 24. We analysed feasibility criteria 3, 9, and 10 by the following procedures:

Calculating the sample size of a future randomised controlled trial (feasibility criterion 3) was based on the primary outcome, namely the immediate transfer quantified with the $MA2_{fluency}$ data. We used the formula:

$$n = \frac{(Z_\alpha + Z_\beta)^2 \times 2\sigma^2}{(\mu_1 - \mu_2)^2}$$

Where n=sample size, Z_α=standard normal z-value for a significance level α=0.05, which is 1.96, and Z_β=standard normal z-value for the power of 80%, which is 0.84 [34]. The pooled standard deviation of both practice groups is named σ; μ_1 and μ_2 are the mean pre-post-intervention difference of the intervention group 1 and intervention group 2, respectively.

The internal responsiveness of the outcome measures (feasibility criterion 9) was determined by calculating the standardised response mean (SRM) for $MA2_{fluency}$ and nP_{norm} between the two assessment time points right before the first and immediately after the last practice session [31]. We performed the analysis for the whole sample, and for the blocked and the random order group separately. The SRM was calculated according to the following formula:

$$SRM = \frac{X_{change}}{SD_{change}} \quad [32]$$

Where X_{change} is the mean change score and SD_{change} the standard deviation of the change scores [31,32]. External responsiveness was determined by calculating a Pearson's correlation coefficient (*r*) between the pre-post changes of the $MA2_{fluency}$ and the ChARMin parameter nP_{norm} [31]. We would expect a strong negative correlation between the changes in the two measures. In case the data were distributed not normally, we calculated the Spearman correlation.

To investigate the influence of other therapies potentially affecting the outcome (feasibility criterion 10), we used data from the whole sample and determined differences in $MA2_{fluency}$ scores between the two assessment blocks performed in week 1. We used a paired T-test (or a Wilcoxon signed rank test in case of not normally distributed data) and calculated effect sizes (ES) according to the following formula (according to [33]):

$$ES = \frac{X_{change}}{SD_{pooled}} \quad [32]$$

Where X_{change} is the mean change and SD_{pooled} the pooled standard deviation. This ES has been recommended for the use of comparisons with dependent groups of the same sample [32].

We interpreted the ES and the SRM according to Cohen's benchmarks: < 0.2: trivial effect, 0.2- 0.5: small effect, 0.5- 0.8: moderate effect, > 0.8: large effect [33]. These thresholds have been described to overestimate the SRM, yet, as there is no other suitable index for SRM, and since ES and SRM obtained from the chosen formula are comparable [32], we used these thresholds also for SRM.

RESULTS

1. Feasibility of inclusion criteria: During familiarisation, two participants were not able to move their more affected arm in the ChARMin robot sufficiently well to play the exergames. One of them had problems moving the arm in the horizontal plane (elbow extension), the other one in both planes.

2. Recruitment rate: Figure 6.4 shows a flowchart displaying the participants' adherence for each appointment with reasons for withdrawal or exclusion. Within one year, 194 children were admitted to our rehabilitation centre (six of them attended two rehab stays). Only 16 children (8.2%) fulfilled the inclusion criteria and could be informed about the study. Three participants refused to take part due to the demands of the study; one person provided no reason. Twelve participants (75% of the eligible children) were recruited. At familiarisation (first appointment), two participants were not able to perform the exergames due to limited upper limb function, one participant was not compliant to attend the following appointments, and one received upper limb Botox injections after the first appointment, which were unexpectedly planned at short notice. Eight participants (66.7% of recruited participants) attended the second appointment. Here, one participant decided to discontinue the study because of the videotaping. As one participant did not attend the last appointment due to illness, six children (50% of recruited participants) completed the whole procedure. Finally, one complete dataset had to be excluded due to technical issues with the robot that led to unreliable data and one participants' data of week 1 had to be excluded from analysis because of the same technical issue, the $MA2_{fluency}$ data due to wrong positioning of the reference point. Therefore, we could obtain complete datasets from 33.3% of the recruited participants and one additional dataset with missing data of the final appointment. Table 6.1 shows the characteristics of the eleven of twelve participants who were

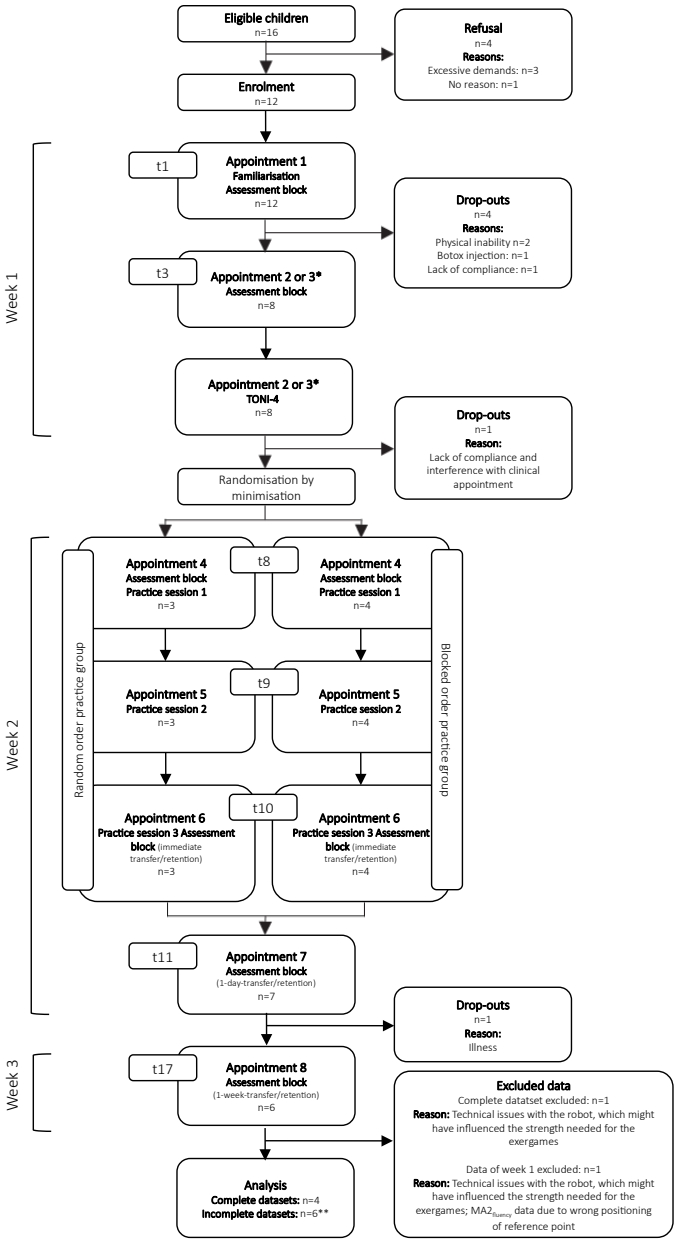


Fig. 6.4. Adherence to the study protocol.
*The Test of Nonverbal Intelligence – Fourth edition (TONI-4) was performed at any day between familiarisation (appointment 1) and randomisation. Hence, the TONI-4 appointment was either the second or the third appointment, the same applies to the second assessment block on t3.
**Participant who dropped out after appointment 7 due to illness was included in the analysis with incomplete dataset. Participant with excluded data of week 1 was incuded in analysis with incomplete dataset.
Abbreviations: t = time point.

Table 6.1. Participant characteristics

Participant number	Age [yrs.]	Sex	Weight [kg]	Height [cm]	Diagnose	MACS level	More affected side	Trained side	Practice group
01	11.7	Male	42.0	146.5	Bilateral spastic cerebral palsy	I	Left	Left	Blocked
02	16.2	Female	86.7	161.5	Unilateral spastic cerebral palsy	I	Left	Left	Blocked
03	10.6	Female	43.0	135.0	Bilateral spastic cerebral palsy	III	Right	Right	Random
04	17.1	Male	60.4	188,5	Myelination disorder, ataxic movement dirsorder	II	Both similar	Left	Random *
05	16.8	Female	59.2	167.0	Traumatic brain injury: bilateral, cerebral movement disorder, arms more accentuated (reduced strength and coordination right upper limb)	I	Right	Right	Random
06	16.9	Female	67.1	170.5	Functional hemispherectomy (Rasmussen encephalitis): hemiparesis, arm more accentuated	III	Left	Left	Not allocated **
07	14.6	Male	31.9	140.5	Spastic cerebral palsy, legs and right side more accentuated	NA	Right	Right	Not allocated ***
08	16.8	Female	16.8	57.8	Unilateral spastic cerebral palsy	II	Left	Left	Not allocated ****
09	6.7	Female	17.2	113.3	Bilateral spastic cerebral palsy	I	Right	Right	Blocked
11	16.2	Male	60.6	164.0	Bilateral spastic cerebral palsy	III	Left	Left	Not allocated **
12	15.1	Male	54.1	168.1	Unilateral spastic cerebral palsy	III	Left	Left	Blocked

*Excluded after the whole study procedure as data might have been distorted due to technical error
**Excluded after familiarisation: physically not able to play the exergame
***Excluded after first assessment block (week1) self-reported compliance not given
****Excluded after first assessment block: Botox injection after the first appointment
One participant (number 10) was not included in the table due to withdrawal of the study including deletion of data.
Abbreviations: yrs=years, kg=kilogramme, cm=centimetre, MACS=Manual Ability Classification System, SD=standard deviation, NA=not applicable.

enrolled in the study, one participant withdrew from the study during week 1, the corresponding data were not published.

3. Sample size calculation for a future randomised controlled trial: Based on the results of the primary outcome ($MA2_{fluency}$ mean difference ± standard deviation: blocked order group 0.01 ± 0.30 points, random order group 0.00 ± 0.20 points), we estimated that 15'355 participants should be included when the $MA2_{fluency}$ would be the primary outcome measure. If we would have selected nP_{norm} as the primary outcome measure (nP_{norm} mean difference ± standard deviation: blocked order group -0.09 ± 0.09 normalised speed peaks, random order group -0.14 ± 0.22 normalised speed peaks), 203 participants would be required.

4. Procedure feasible for participants: For 50% of the included participants, the procedure was not feasible. Three participants were not able to complete 30 trials during the three practice sessions (instead of 30/30/30, the number of trials were 20/20/14, 24/24/24, and 28/23/20). Their practice sessions had to be shortened. We paid attention that the sessions did not take longer than two hours because the burden for the participants would have been too high. That was one reason for aborting the sessions. With time advancing in the practice sessions, we generally noticed a decrease of motivation and concentration and self-reported fatigue, which were also reasons for abortion. One participant prematurely stopped the session due to pain in the shoulder joint.

5. The scheduling procedure proved difficult, as the appointments needed to be planned during the regular inpatient rehabilitation stay. As other therapies had to be considered, it was difficult and time-consuming to plan all the appointments on the same half day. One participant's schedule had to be postponed one week due to illness of the participant during the first measurement week. As it was not possible to schedule the follow-up appointment due to the participant's discharge, the last assessment session was not performed. Furthermore, blinding of the assessors was difficult, because some assessors who assessed the primary outcome measure worked as therapists in the same room and might have noticed how children were practising.

6. The randomisation procedure worked well. Despite the occasional absence of the person primarily in charge of the randomisation by minimisation process, the procedures were always completed in time. In one case, a third person (also not involved in the study) had to operate the randomisation by minimisation program via telephone instruction by the study nurse, as both the study nurse and her deputy were unavailable.

7. Handling of the robot: One complete dataset and the data of the first week's appointment of one participant had to be excluded due to a technical issue (weight-support rope was not properly running over a deflection pulley). While

this was not a safety issue, it might have influenced the performance during the exergames as there was an increased friction in the movement directions of shoulder flexion and extension. A total of nine single trials had to be repeated due to errors (e.g. when the participant moved too fast, the robot stopped due to safety reason). Finally, in some few specific configurations, the robot started oscillating, thus reducing the accuracy of the recorded data. The oscillations occur as an interaction between the patient and the robot, i.e. the movement or oscillation of the patient's arm are amplified by the active joint friction compensation of the robot, leading to an overshoot of the movement. While this occurred in some of the sessions of single patients, the reliable determination of the number of these oscillations was not possible.

8. Amount of data: The therapist assessing the $MA2_{fluency}$ videos had to evaluate a total of 35 videos within one year. Analysing all six videos of one participant costed approximately 35 minutes. In addition, calculating the parameters of the ChARMin data in Matlab from one practice session (i.e. 30 log files) took approximately 43 minutes. For three practice sessions, this amounted to more than two hours. To analyse the exergame test data (6 x 6 log files), approximately 50 minutes were needed. In case of well-functioning software without any program errors, approximately three hours of data processing were needed per participant. Additional time was required occasionally when program errors occurred (IDs 01, 02 and 03 without errors; ID 04 two errors; IDs 05 and 06 could initially not be processed at all and two software updates were needed to fix the bugs). Overall, approximately 6.5 hours per patient were required to analyse the data.

9. Responsiveness of outcome measures: $MA2_{fluency}$ sum-scores per participant are shown before, immediately after, one day, and one week after the last practice session in Figure 6.5. At the immediate transfer, which was our primary outcome, three participants improved (by four, two, and two points), one remained stable, and two participants deteriorated (by minus one and minus four points). Internal responsiveness of the $MA2_{fluency}$ showed a trivial effect for the whole group and the blocked order group (SRM = 0.18 and 0.15, respectively) and a small effect for the random order group (SRM = 0.24). The internal responsiveness analyses of the nP_{norm} data showed that these were moderate to large effects (SRM =-0.84, -0.95, and-0.62 for the whole group, the blocked order group, and the random order group, respectively). External responsiveness: the change in $MA2_{fluency}$ correlated weakly and non-significantly with the change in nP_{norm} for the whole group ($r = -0.25$, $p = 0.63$) and very weak for the blocked group ($r = 0.03$, $p = 0.97$). As in the random group only two datasets were available, a correlation could not be calculated.

10. Parallel therapies potentially affecting the outcome: Figure 6.6 shows the distribution of the different therapies during week 1 (between the two assessment blocks) and during week 2 (between the assessment block before and the one immediately after the practice phase). For one patient, the exergame test as well as the $MA2_{fluency}$ data were not included because of the previously mentioned hardware issue and a mistake in the measurement taken during the $MA2_{fluency}$ (the reference point where the objects are placed during the test were not equally positioned at both time points). The paired samples T-test between the two assessment time points showed trivial ES and no significant changes neither of the $MA2_{fluency}$ (ES =-0.10, t =-0.74, p = 0.50) nor the nP_{norm} data (ES-0.08, t = -0.49, p = 0.65).

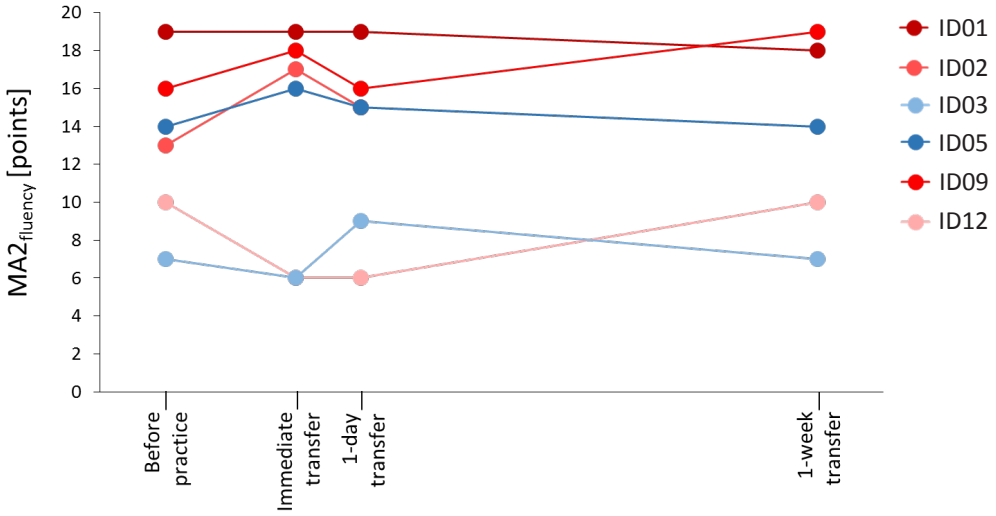


Fig. 6.5. Immediate, 1-day, and 1-week transfer. Displayed are the sum scores of the Melbourne Assessment 2, subscale "fluency" ($MA2_{fluency}$) for each participant at the time points before, immediately after, one day and one week after practice. Participants are represented with different shades of colours, reddish colours represent blocked practice order, blueish colours represent random practice order.

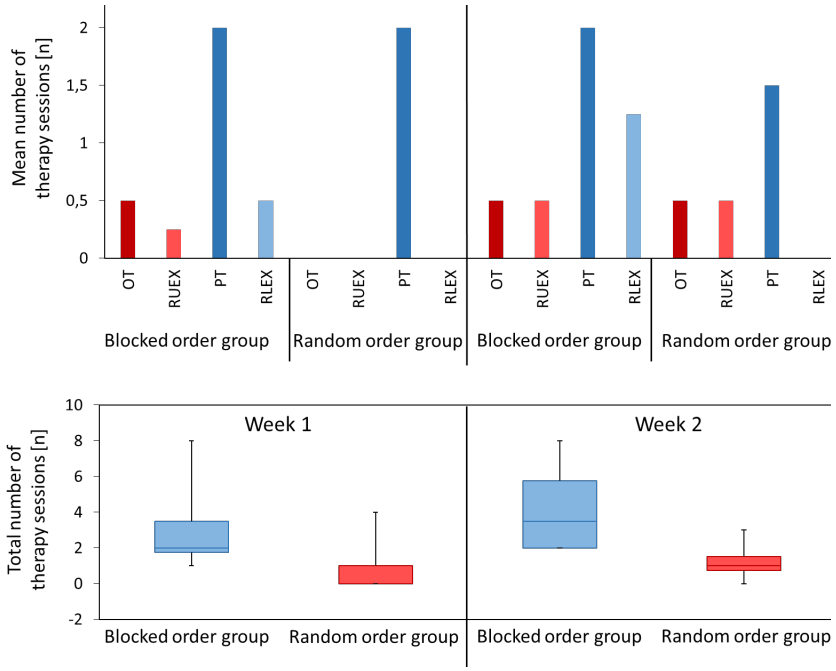


Fig. 6.6. Additional therapies during week 1 and week 2. A) The mean numbers of the different therapies are displayed for both practice groups and both weeks. B) Boxplots of the total number of therapies that have taken place during week 1 and week 2 for both practice groups.
Abbreviations: n=number, OT=Occupational Therapy, RUEX=Robotics Upper Extremities, PT=Physiotherapy, RLEX=Robotics Lower Extremities.

DISCUSSION

With this pilot study, we investigated the feasibility of an RCT assessing blocked and random practice order of two versions of an exergame with the ChARMin exoskeleton to improve upper limb functions in children with neuromotor impairments. Of the ten predefined feasibility criteria, we considered six criteria to be partially feasible: in-/exclusion criteria, scheduling, randomisation procedure, handling of the robot, amount of data, influence due to parallel therapies; and four criteria to be not feasible: recruitment rate, participants' procedure, responsiveness of primary outcome measure, and sample size.

Concerning the feasibility of inclusion criteria, we had selected the MACS level as indicator for the ability to move the upper extremity. For the current study, the MACS proved not suitable as an inclusion criterion, as two participants were not able to perform the movements required for practicing the exergames with ChARMin. It follows that the inclusion criteria were not sensitive enough and, therefore, this criterion is only partially feasible. For a future trial, we would recommend to include

strength tests. Especially the shoulder abductors and elbow extensors should have at least a manual muscle score of 3 (movement against gravity over the whole range of motion) [35]. Two participants would have needed support of ChARMin to perform the exergames in the frontal plane against gravity. Yet, we had decided a priori to refrain from including support by the robot due to standardisation of the interventions. However, ChARMin was designed to train and evaluate also more severely affected patients and the level of weight- and movement support can be adjusted, also for children with more severely affected upper limb abilities [27]. So, another option would be to develop and include a standardised procedure for defining the level of weight support and movement support for patients with different upper limb ability levels and to adapt the support by the robot accordingly to each individual patient.

An eligibility-recruitment rate of 75% and a recruitment and completion of procedure-proportion of 50% can be considered low. With these rates, recruiting would take approximately 1280 years to achieve 15'355 participants (primary outcome measure: $MA2_{fluency}$) or 17 years (nP_{norm} as primary outcome). As a comparison, in a study protocol evaluating the dosage and type of constraint-induced movement therapy in children with cerebral palsy, an eligibility-recruitment rate of 90% is estimated [36]. Another study assessing the efficacy of modified constraint-induced movement therapy also in children with cerebral palsy, screened 76 potential participants and could include 52 of them [37]. Fifty of them attended all the measurements and, accordingly, the recruitment and analysable-dataset proportion was 96% [37]. Different recruitment strategies have been evaluated, yet, it has been stated that there is a lack of knowledge about effective recruitment strategies [38]. In our study, any other strategy would have led to the same outcome as we screened and recruited all eligible participants attending our rehabilitation centre. Unfortunately, since the ChARMin robot is not a commercially available device, and a multicentre trial to increase the number of eligible participants would not be possible. A randomised controlled trial requiring these sample sizes is practically impossible within our setting.

The sample size calculation based on the $MA2_{fluency}$ is not feasible to achieve. Calculations based on the exergame test data were lower, yet, also not feasible to reach with one device in one centre. It has been recommended to interpret sample size calculations based on pilot studies with caution as estimated treatment effects might be biased due to the low sample size in the pilot study [19]. In fact, estimates of ES based on data obtained from pilot studies have been shown to be insufficiently accurate to be used for the decision-making about whether or not a major trial should be funded [39]. They should, therefore, never be the only criterion to decide about funding of a trial.

Concerning the feasibility of the study procedure from a participants' point of view, we had to adjust the procedure several times. The schedule had to be shortened for three participants due to fatigue, low motivation, and in one case due to shoulder

pain. Especially in such a heterogeneous study population, it is a challenge to find the number of practice trials sufficient to induce motor learning while being feasible for all participants at the same time. The length and amount of trials of the practice sessions should be individualised according to the participants' abilities. Yet, in a randomised controlled trial, individually tailoring the interventions could cause an imbalance in terms of numbers of repetitions. One solution for a future trial would be to increase the number of sessions while simultaneously reducing the number of trials per session. Another approach would be to perform multiple N-of-1 trials, which is indicated when there is substantial uncertainty about the effectiveness of comparable interventions and allows for more individual interventions [40]. Since every participant would serve as their own control, they would practice under both practice orders in a randomised order. Yet, as a motor learning effect would probably be counted as a lasting carry-over effect, a washout period of sufficient length between the two practice periods needs to be considered [40]. This would make it difficult to implement such a design in our setting, where the patients' length of stay is limited. Hence, practice sessions at home with mobile devices should be discussed for future research.

Scheduling was difficult mainly due to the appointments of the regular rehabilitation schedule, which were treated with higher priority. The rigidity of the study procedure (practice sessions on three consecutive days, follow-up appointments exactly one day and one week after the last practice session, etc.) negatively affects the possibilities of rescheduling if a participant had missed an appointment. A simpler, more flexible study design would make the (re-) scheduling easier but might affect the learning process in a negative way if conditions would change continuously.

The randomisation procedure worked well, especially since several people were involved in the procedure, absences could be bridged with assigned deputies. Yet, it is recommended to advise a third person a priori to substitute for the deputy in case of his or her absence. A three-person randomisation team would be feasible in our setting.

Technical issues of the current version of ChARMin make the handling of the robot only partially feasible. As the technology is relatively novel, ChARMin may need to undergo some further developments (e.g. preventing the occurrence of movement oscillations or ensuring that the weight-support-rope is running properly) before implementing it in a future trial.

Concerning the evaluation of the outcome measures, evaluating the $MA2_{fluency}$ videos was feasible. Yet, processing the ChARMin data (exergame tests and practice) was time consuming. A more automated program would be needed to handle larger amounts of data.

The $MA2_{fluency}$ was none to moderately responsive when applied as primary outcome in our study. When looking at the performance of the single participants, after the last practice session, three of them had improved beyond the minimal detectable change (MDC) of 1.84 points for the $MA2_{fluency}$ [29] compared to the onset of practice. One participant did not show a change and two participants deteriorated, one of them beyond the MDC. In comparison, children with cerebral palsy showed a mean improvement of 0.97 points in the $MA2_{fluency}$ after an extensive eight-weeks upper limb training [29]. That same study obtained an SRM of 1.84 (highly responsive) for the $MA2_{fluency}$ [29]. Yet, the intervention was 4.5h hours per week of uni- or bilateral upper limb training for eight weeks [29]. Such an intensive intervention is more likely to yield a pre-post-intervention effect compared to our protocol. Indeed, responsiveness of an outcome measure is determined in a group of patients who have made a true change [41]. As we could not analyse changes due to the intervention in our study, this cannot be verified. The external responsiveness analysis showed a weak correlation between the change measured with $MA2_{fluency}$ and the change measured with nP_{norm} for the whole group. This means that the two measures did not reflect the same construct of measuring change. The internal responsiveness analysis of the nP_{norm} data (all participants included) showed moderate to large SRMs, while the $MA2_{fluency}$ data provided trivial SRMs. The internal responsiveness nicely shows why the correlation of the changes measured with the two measures was weak.

Using a robotic device enables accurate, objective, and sensitive assessment of body functions of children [42]. From this point of view, choosing a robotic device to measure the primary outcome might be recommended. Nevertheless, as we were interested in the transfer and we wanted to measure fluency, we measured fluency outside the robotic device. Therefore, we chose $MA2_{fluency}$ as the primary outcome measure.

Concerning the influence of the parallel therapies, we did not find a change in the $MA2_{fluency}$ between the two time points at week 1. As even the nP_{norm} ES was trivial, in contrast to the large SRMs found during the practice week, it is likely that parallel therapies indeed did not change movement fluency.

Limitations

There were some limitations in this study. We performed a pilot study in our rehabilitation centre using this specific robotic device. While several feasibility issues can be generalised to other centres, for example, concerning patient recruitment or planning of inpatient appointments, other issues might be specific for our setting, for example, the specific robotic device and its output, restricting generalisability.

We chose to calculate the sample size of a future trial with the primary outcome data. However, determining the sample size based on a meaningful effect such as quantified by the minimal clinically important difference would result in a more

appropriate power calculation, including the magnitude of improvement and how patients value the change.

Applying a standardised questionnaire to record how the participants felt (motivation, fatigue, and enjoyment) would have increased data quality and revealed specific descriptors how the participants experienced the study procedure. The Physical Activity Enjoyment Scale Questionnaire would be one example of such a questionnaire [43,44]. Yet, applying an additional outcome measure would have prolonged the assessment protocol even more.

There were also limitations in the study design concerning motor learning aspects. The primary outcome assessor was sometimes present during study appointments. This was unavoidable, as he also worked as a therapist in the same room with other patients. Yet, we tried to keep the ratings blinded by making the period between the appointments and the video-analysis several weeks long and censoring the videos for the assessment time point. Therefore, we consider this issue as marginal.

Another motor learning issue might have been an insufficient contrast between the interventions (blocked versus random order). The rationale behind the choice of these two interventions is based on the hypothesis that variations of tasks based on different motor programs result in a contextual interference effect [5]. Different spatial configurations (as it is the case in the current pilot study by moving in different planes) have been suggested to require different motor programs [4]. This feature has also been chosen in a study evaluating throwing tasks on a horizontal and a vertical target in six-year-old children [45]. Results showed a superiority of random over blocked practice when comparing the total score during acquisition, at retention, and transfer [45]. However, in addition to the target positions varying in space, the projectiles (four different kinds of balls) were also varied [45]. In our study, the spatial characteristics of the exergames was the only varying factor determining contextual interference. As variations coming from more than one parameter could increase the contextual interference effect [5], we would suggest to increase the contrast between the two study arms for a future trial.

Finally, we could only include six participants in our study, which limits us in making generalisations concerning the motor learning outcomes.

CONCLUSION

We could show that it is not feasible to perform a large RCT when using the design as evaluated in this pilot study. The main reasons are the low recruitment rate, the demanding study procedures, and the large sample size that would be needed for a main trial. We made several suggestions how to improve the study design and discussed alternatives such as n-of-1 trials.

Studies on whether the contextual interference effect in children with neuromotor disorders exist are still needed, as this could guide future research and clinical treatment of this vulnerable patient group.

ACKNOWLEDGEMENTS

We would like to thank all the participants and their parents for contributing to this study. Furthermore, we would like to thank Jan Lieber, Nicole Iten, Anne-Laure Mouton and Jeffrey Keller of the Research Department of the Swiss Children's Rehab for their support during the study.

AUTHOR'S CONTRIBUTIONS

JVG conceptualised and designed the study, conducted measurements and interventions, and drafted the manuscript. CHGB supported the conceptualisation of the study and substantively revised the manuscript. AG conducted measurements and interventions and revised the manuscript. UK created the new software that is described in this study, supported the conceptualisation of the study, and revised the manuscript. HJAvH supported the conceptualisation of the study, substantively revised the manuscript, and obtained the funding.

All authors have approved the submitted version and have agreed both to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which they were not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

FUNDING

This study protocol has been supported by the Clinical Research Priority Programme, Neuro-Rehabilitation of the University of Zurich, and the Mäxi Foundation.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study has been approved as part of the ChARMin project (sub-project 5: Motor learning) by the Ethics Committee of the Canton Zurich (BASEC-Nr. PB_2016-02450) and the Swiss Agency for Therapeutic Products (Swissmedic reference number: 2015-MD-0009).

According to the Ethics Committee's guidelines all the participants should give their verbal consent, children of 15 years and older and all the legal representatives should

give informed written consent. All the participants and their legal representatives have the right to withdraw their participation at any time.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AVAILABILITY OF DATA AND MATERIAL

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

ADDITIONAL FILE 6.1

Contextual interference in children with brain lesions: a pilot study investigating blocked vs. random practice order of an upper limb robotic exergame

Judith V Graser, Caroline HG Bastiaenen, Anja Gut, Urs Keller, Hubertus JA van Hedel

Observations of motor learning

As there were only six datasets available, we present some aspects about motor learning worth discussing in an exploratory manner in this supplementary file.

Transfer

A visualisation of the Melbourne Assessment 2, subscale Fluency ($MA2_{fluency}$) sum-scores per participant at the time points before, immediately after, one day and one week after the last practice session is displayed in Figure A6.1 (equal to Figure 6.5 in the main manuscript).

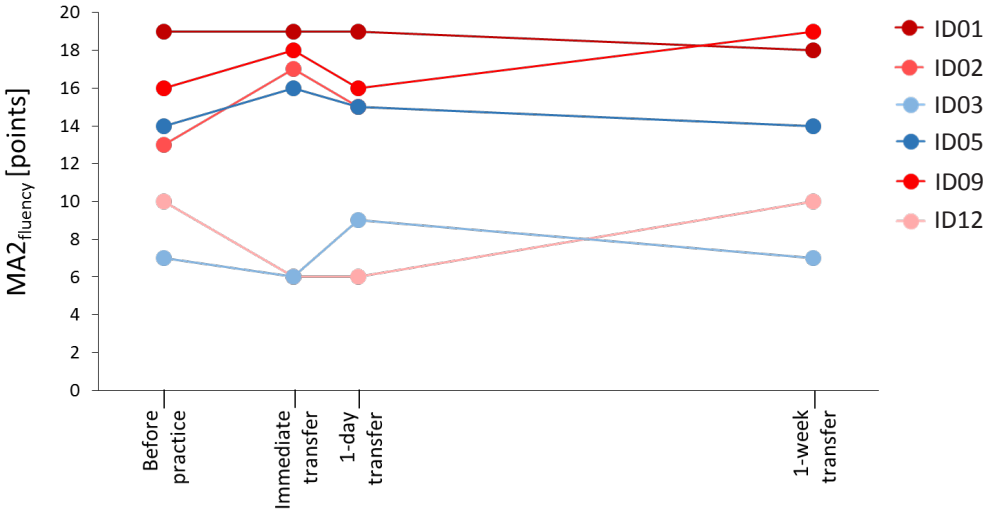


Fig. A6.1. Immediate, 1-day and 1-week transfer. The Melbourne Assessment 2 subscale Fluency sum scores ($MA2_{fluency}$; Y-axis) are plotted for each time point (X-axis) for each participant. Participants are represented with different shades of colours, reddish colours represent blocked practice order, blueish colours represent random practice order. (This Figure is equal to Figure 6.5 in the main manuscript).

At the immediate transfer, which we initially chose as primary outcome time point, three participants improved beyond the minimal detectable change (MDC) of 1.84 points for $MA2_{fluency}$ [29]. However, one participant did not show a change and two participants deteriorated, one of them beyond the MDC. It is noteworthy, that at least some of the participants showed a substantial change in performance of

the $MA2_{fluency}$ within such a short time. In comparison, children with cerebral palsy showed a mean improvement of 0.97 points in the $MA2_{fluency}$ after an extensive eight-weeks upper limb training [29].

Retention

Data obtained during the exergame tests and used for retention are displayed in Figure A6.2.

At immediate retention, four participants improved (-4.3, -3.3, -7.3 and -2.0 velocity peaks per movement, normalised to the actual distance). The other two showed slight deteriorations (0.5 and 0.9 velocity peaks per movement, normalised to the actual distance). These two participants started with lower performance levels (i.e. higher numbers of peaks) compared to the four participants who improved during practice. Yet, they improved at delayed retention.

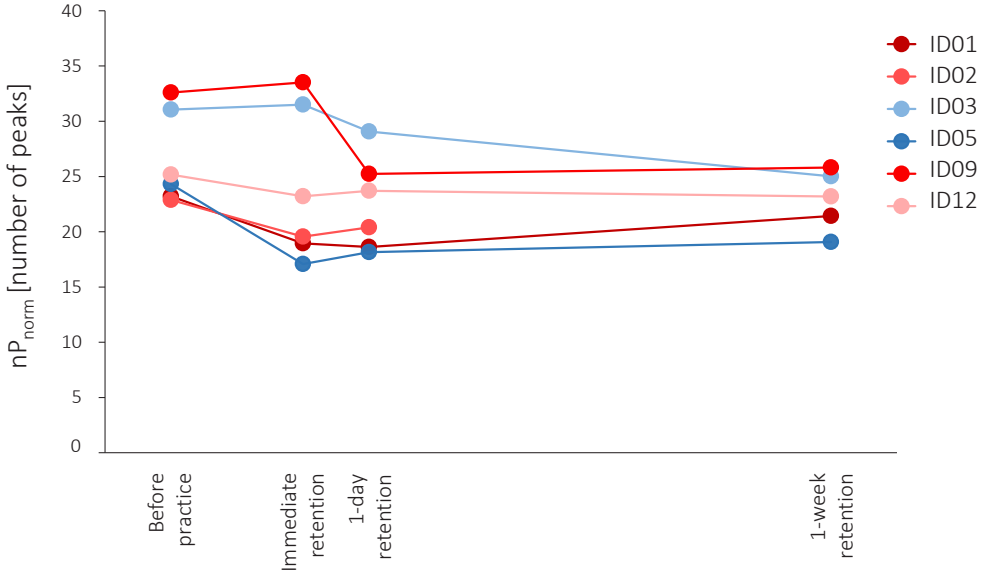


Fig. A6.2. Immediate, 1-day, and 1-week retention. The mean numbers of velocity peaks normalised to the actual distance (nP_{norm}) over one game test, i.e. three trials on the horizontal and three trials on the frontal plane (Y-axis) are plotted for each time point (X-axis) for each participant. Participants are represented with different shades of colours, reddish colours represent blocked practice order, blueish colours represent random practice order.

These 'groups' were not reflecting the practice groups. We assume that the slight deteriorations in the two participants might reflect normal trial-to-trial variability or they might have temporarily decreased their performance at the end of the practice period due to, for example, fatigue. Indeed, at the one-day retention, their performance improved, which^we could explain as a sort of recovery (e.g. from

fatigue). This could be an indication that retention test should be conducted after a break to allow participants to recover from the intensive practice.

When comparing the data at one-week retention with the data from before the practice sessions, all participants (with the exception of one participant who did not attend the one-week retention test) improved their performance by -1.8 to -6.8 velocity peaks.

Acquisition

Figure A6.3 shows the plotted data of all trials and participants with a linear trend-line over the course of the three practice sessions. No general trend can be observed. Some of the participants seem to deteriorate over the three sessions while some seem to improve. Some participants show fluctuations during or between the sessions, while others perform on a relatively steady level.

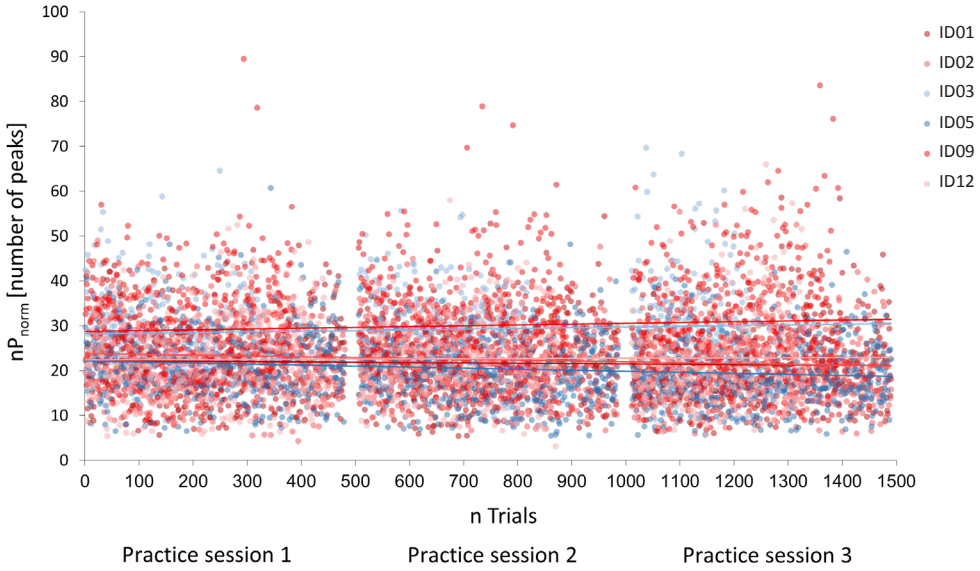


Fig. A6.3. Acquisition: performance during the practice phase. The number of velocity peaks normalised to the actual distance (nP_{norm}; Y-axis) are plotted (participants are represented with different shades of colours, reddish colours represent blocked practice order, blueish colours represent random practice order.) with linear trend lines for each participant (according opaque colours) over all trials of all three practice sessions (X-axis).

This widespread fluctuation in performance during acquisition within each practice session and each individual participant is eye catching. There are some general factors like motivation [46] or fatigue [47] influencing variability in motor performance and learning. In our study, we identified also some movement-related factors. Firstly, within each trial, participants head for eight targets and move each time back to the

centre point meaning that all 16 movements point in another direction. Depending on a child's motor impairments, movement experience, and preferences, one movement direction might be easier to be performed fluently compared to another one. Secondly, the participants might have preferences for moving in the horizontal or frontal plane. Such aspects could be partly responsible for the widespread performance within each participant.

REFERENCES

1. Krakauer JW. Motor learning: Its relevance to stroke recovery and neurorehabilitation. Curr Opin Neurol. 2006;19(1):84–90.
2. Muratori LM, Lamberg EM, Quinn L, Duff S V. Applying principles of motor learning and control to upper extremity rehabilitation. J Hand Ther. 2013;26(2):94–103.
3. Schmidt RA, Lee TD. Motor learning and performance. From principles to application. 5th editio. Champaign; 2014.
4. Bernstein NA. The co-ordination and regulation of movements. Oxford: Pergamon Press; 1967.
5. Magill RA, Hall KG. A review of the contextual interference effect in motor skill acquisition. Hum Mov Sci. 1990;9(3–5):241–89.
6. Shea JB, Morgan RL. Contextual interference effects on the acquisition, retention, and transfer of a motor skill. J Exp Psychol Hum Learn Mem. 1979;5(2):179–87.
7. Shea JB, Zimny S. Context effects in memory and learning movement information. In: Magill, Richard A, editor. Memory and control of action. Amsterdam: North-Holland; 1983. p. 345–66.
8. Lee TD, Magill RA. Can forgetting facilitate skill acquisition? In: Goodman D, Wilberg R, Franks I, editors. Differing perspectives in motor learning, memory, and control. New York: Elsevier Science; 1985. p. 3–21.
9. Poto CC. How forgetting facilitates remembering: An analysis of the contextual interference effect in motor learning. 1988.
10. Thürer B, Gedemer S, Focke A, Stein T. Contextual interference effect is independent of retroactive inhibition but variable practice is not always beneficial. Front Hum Neurosci. 2019;13.
11. Shewokis PA, Del Rey P, Simpson KJ. A test of retroactive inhibition as an explanation of contextual interference. Res Q Exerc Sport. 1998;69(1):70–4.
12. Brady F. Contextual interference: a meta-analytic study. Percept Mot Skills. 2004;99:116–26.
13. Graser JV, Bastiaenen CHG, van Hedel HJA. The role of the practice order: A systematic review about contextual interference in children. PLOS One. 2019;14(1).
14. Aurich-Schuler T, van Hedel HJA, Labruyère R. Roboterunterstützte Lokomotionstherapie bei Kindern in der Neuroreha. neuroreha. 2018;10(03):119–26.

15. van Hedel HJA, Lieber J, Ricklin S, Meyer-Heim A. Die praktische Anwendung von Exergames und virtueller Realität in der pädiatrischen Rehabilitation. neuroreha. 2017;09(01):35–40.

16. Gerber CN, Kunz B, van Hedel HJA. Preparing a neuropediatric upper limb exergame rehabilitation system for home-use: A feasibility study. J Neuroeng Rehabil. Journal of NeuroEngineering and Rehabilitation; 2016;13(33):1–12.

17. Duret C, Grosmaire A-G, Krebs HI. Robot-assisted therapy in upper extremity hemiparesis: Overview of an evidence-based approach. Front Neurol. 2019;10:1–8.

18. Prado MTA, Gonçalves Luiz Fernani DC, Dias da Silva T, Smorenburg ARP, de Abreu LC, Bandeira de Mello Monteiro C. Motor learning paradigm and contextual interference in manual computer tasks in indivisuals with cerebral palsy. Res Dev Disabil. 2017;64:56–63.

19. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: The what, why and how. BMC Med Res Methodol. 2010;10(1).

20. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. BMJ. 2016;355(i5239).

21. Chan A-W, Tetzlaff JM, Altman DG. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7.

22. Graser JV, Bastiaenen CHG, Keller U, van Hedel HJA. Contextual interference in children with brain lesions: protocol of a pilot study investigating blocked vs. random practice order of an upper limb robotic exergame. BMC Pilot Feasibilty Stud. 2020;6(156).

23. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Ohrvall A-M, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: Scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006;48(7):549–54.

24. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther. 1987 Feb;67(2):206–7.

25. Altman DG, Bland JM. Treatment allocation by minimisation. Br Med J. 2005;330(April):843.

26. Ritter N, Kilinc E, Navruz B, Bae Y. Test Review: L. Brown, R. J. Sherbenou, & S. K. Johnsen 'Test of Nonverbal Intelligence-4' (Toni-4). Austin, TX--PRO-ED, 2010. J Psychoeduc Assess. 2011;29(5):484–8.

27. Keller U, Van Hedel HJA, Klamroth-Marganska V, Riener R. ChARMin: The first actuated exoskeleton robot for pediatric arm rehabilitation. IEEE/ASME Trans Mechatronics. 2016;21(5):2201–13.

28. Randall M, Johnson L, Reddihough D. The Melbourne Assessment 2 [Internet]. [cited 2019 Mar 29]. Available from: https://www.rch.org.au/melbourneassessment/

29. Wang T-N, Liang K-J, Liu Y-C, Shieh J-Y, Chen H-L. Psychometric and clinimetric properties of the Melbourne Assessment 2 in children with cerebral palsy. Arch Phys Med Rehabil. 2017;98(9):1836–41.

30. Graser JV, Prospero L, Liesch M, Keller U, van Hedel HJ. Reliability of and practical lessons learned from robotic upper limb assessments in children undergoing neurorehabilitation. Submitted.

31. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol. 2000;53(5):459–68.

32. Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integr Care. 2002;2.

33. Cohen J. Statistical power analysis for the behavioral sciences. 2nd editio. Hillsdale: Lawrence Erlbaum Associates; 1988.

34. Hickey GL, Grant SW, Dunning J, Siepe M. Statistical primer: Sample size and power calculations — why, when and how? Eur J Cardio-Thoracic Surg. 2018;54:4–9.

35. Kendall McCreary E, Peterson Kendall F, Geise Provance P. Muscles: Testing and function. Philadelphia: Lippincott Williams and Wilkins; 1993.

36. Landesman Ramey S, DeLuca S, Stevenson RD, Case-Smith J, Darragh A, Conaway M. Children with Hemiparesis Arm and Movement Project (CHAMP): Protocol for a multisite comparative efficacy trial of paediatric constraint-induced movement therapy (CIMT) testing effects of dosage and type of constraint for children with hemiparetic cerebra. BMJ Open. 2019;9(1).

37. Aarts PB, Jongerius PH, Geerdink YA, van Limbeek J, Geurts AC. Effectiveness of modified constraint-induced movement therapy in children with unilateral spastic cerebral palsy: A randomized controlled trial. Neurorehabil Neural Repair. 2010;24(6):509–18.

38. Treweek S, Mitchell E, Pitkethly M, Cook J, Kjeldstrøm M, Johansen M, et al. Strategies to improve recruitment to randomised controlled trials (Review). Cochrane Database Syst Rev. 2010;1(MR000013).

39. Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch Gen Psychiatry. 2006;63(5):484–9.

40. Kravitz R, Naihua D, Sunita V, Jiang L, The DEcIDE Methods Center N-0f-1 Guidance Panel. Introduction to N-of-1 trials: Indications and barriers. In: RL K, Duan N, editors. Design and implementation of N-of-1 trials: a user's guide. Rockville: Agency for Healthcare Research and Quality; 2014. p. 1–11.

41. Carter R, Lubinsky J, Domholdt E. Rehabilitation research: Principles and applications. 4th editio. St. Louis: Elsevier Saunders; 2005.

42. Gilliaux M, Lejeune TM, Detrembleur C, Sapin J, Dehez B, Selves C, et al. Using the robotic device REAplan as a valid, reliable, and sensitive tool to quantify upper limb impairments in stroke patients. J Rehabil Med. 2014;46(2):117–25.

43. Latorre Román PÁ, García Pinillos F, Navarro Martínez AV, Izquierdo Rus T. Validity and reliability of Physical Activity Enjoyment Scale questionnaire (PACES) in children with asthma. J Asthma. 2014;51(6):633–8.

44. Boffoli N, Foley JT, Gasperetti B, Yang SP, Lieberman L. Enjoyment levels of youth with visual impairments playing different exergames. Insight Res Pract Vis Impair Blind. 2011;4(4):171–6.

45. Granda Vera J, Montilla MM. Practice schedule and acquisition, retention, and transfer of a throwing task in 6-yr.-old children. Percept Mot Skills. 2003;96(3 Pt 1):1015–24.

46 Wulf G, Lewthwaite R. Optimizing Performance through Intrinsic Motivation and Attention for Learning: The OPTIMAL theory of motor learning. Psychon Bull Rev. 2016;22(6):1–35.

47. Branscheidt M, Kassavetis P, Anaya M, Rogers D, Huang HD, Lindquist MA, et al. Fatigue induces long-lasting detrimental changes in motor-skill learning. Elife. 2019;8:e40578:1–25.

CHAPTER 7

GENERAL DISCUSSION

CHAPTER 7

"YOU CAN ALWAYS CHANGE YOU PLAN BUT ONLY IF YOU HAVE ONE."
Randy Pausch, the last lecture

INTRODUCTION

At the time of planning this doctoral thesis, the overall objective was to investigate blocked versus random practice order of variations of a motor task in children with brain lesions in a randomised controlled trial. We aimed to gain knowledge about the effectiveness of these two practice orders. We had an initial plan but along the way we adapted and changed it based on the knowledge we gained with every step. Experience led to these changes. Yet, changing of plans was only possible because we had this initial plan. In the end, therefore, we did not conduct a main randomised controlled trial, but we assessed the feasibility of a randomised controlled trial, a step that was not part of the plan in the beginning. The results of this pilot study changed the perspective on evaluating effectiveness with a large randomised controlled trial to quite some extent. With the current knowledge, it is not advisable to perform a main randomised controlled trial within the given context. This chapter displays the way from the initial idea to recommendations for future research. In our setting, we recommend considering different study designs to evaluate effectiveness. Great potential lies in the inclusion of the children, their families, and caregivers throughout the whole research process.

CONTEXTUAL INTERFERENCE IN HEALTHY CHILDREN AND CHILDREN WITH BRAIN LESIONS

The first part of this PhD thesis dealt with the existing literature about the contextual interference effect (blocked versus random order practice) in healthy children and children with congenital or acquired brain lesions. We included 25 papers for the analysis. The methodological quality of the included studies was generally low and the risk of bias high. For healthy, typically developing children, we found limited to moderately consistent evidence supporting parts of the contextual interference effect (acquisition, retention or transfer) for some of the tasks evaluated in the included studies. Only one of the 25 included papers evaluated children with brain lesions, namely a group of children with cerebral palsy [1]. Yet, unfortunately, results were not presented for this participant group separately. Concluding, for the paediatric neurorehabilitation, evidence is missing. For healthy children, evidence with a sound methodology is lacking.

Heterogeneity of the studies

Initially, we planned to perform a meta-analysis with the data of the included studies. However, in addition to the poor quality of the studies, which would challenge the meta-analytical results, we found that it was impossible to reasonably cluster the included studies accordingly. There were only small numbers of studies that matched to key factors such as similar study designs, learning tasks, length of exercise phases, and number of exercise sessions or outcome measures. When, for example, the practiced task was comparable, the time periods of practice differed a lot between the studies, which made a comparison not meaningful. Therefore, we refrained from performing a meta-analysis and conducted a best evidence synthesis [2].

Several reviews of studies evaluating the contextual interference effect in mainly adult participants have been conducted earlier. Conclusions of these reviews differed: one review concluded that contextual interference is a basic and general feature of learning processes, as it has been observed in many tasks [3]. Others stated that the contextual interference effect was dependent on the task characteristics (e.g. laboratory versus non-laboratory tasks, based on the same or on different motor programmes) and on subject characteristics (e.g. beginner versus experienced learner) [4, 5] and that no universal practice schedule could be accounted as beneficial in all contexts and types of persons [6]. Contextual interference seems to be an effect, which is highly dependent on the circumstances in which motor learning takes place. In view of this, one might have guessed that conducting a methodologically sound randomised controlled trial with generalisable results could be difficult in our field.

Paediatric versus adult population

With lacking evidence in the paediatric field, it might thus be helpful to look beyond paediatrics, into the adult population with neurological diagnoses. As stated in chapter 2 of this dissertation, in the past, studies evaluating contextual interference in adults after stroke reported differing results [7–9]. One very recently published study evaluated random versus blocked practice of spooning movements (different trajectories) [10]. The authors concluded that while both practice orders induced improvement and maintenance of feeding skills, yet, only random order practice led to a positive transfer to the real life situation (eating corn flakes) [10]. This study is an important contribution as it addresses contextual interference in the field of neurorehabilitation and includes an activity of daily life.

In medicine, research results obtained in adults are often translated to the paediatric population, for example, drug prescription [11, 12]. The reason is a lack of evidence in the paediatric population [11, 12]. It is a vulnerable group and more strict regulations need to be considered when performing research with children. This might be one of the reasons contributing to the fact that fewer studies are performed with children. Furthermore, in paediatric neurorehabilitation in a small country like Switzerland,

the number of eligible participants in is limited. This limitation is increased if the inclusion criteria are very specific (e.g. if certain cognitive or motor skills or a high level of compliance are required). Then, it is difficult to recruit an appropriate number of participants within a reasonable period. Yet, it would be important to differentiate between adults and children as children are not just small adults [13, 14]. One can assume that the same is true for motor learning, and there are several reasons to endorse this assertion. For example, human beings constantly adapt to the environment and develop accordingly [15]. Hence, adults have more movement experience than children and move most probably in a different environment. A study, looking at three age groups of a) seven to nine years (representing childhood), b) 19 to 25 years (representing young adulthood), and c) 66 to 80 years (old age), showed that motor performance increases from childhood to young adulthood and decreases again until old age [15]. Findings from neuroimaging studies suggest that the maturation of the motor system is not terminated at late pre-school age [16]. Sensorimotor developmental processes seem to be completed around puberty (e.g. timed performance of repetitive and alternating hand movements, decrease of associated movements), and the more complex the task, the later the performance plateau is reached [17]. Furthermore, it has been proposed that children below 12 years of age do not have the ability to adapt rapidly to changed motor patterns (in this study it was a split-belt treadmill walking task) because of a general-purpose motor memory, which is not yet ready to handle different situations [18]. Additionally, procedural learning has an age effect, probably based on the development of the ability to inhibit inappropriate motor responses [19], which shows an active development period between four and nine years [20]. A study has shown that six-year-olds seem to have more difficulties to adapt their motor responses to the task in the initial learning phase compared to ten-year-olds and adults [19]. Hence, motor and learning requirements are not the same for children, especially at a young age, as they are for adults. Given these differences, and certainly many others, it seems appropriate to consider the paediatric and adult populations separately. Hence, more research is needed, not only for healthy children but especially for children with neuromotor disorders.

Concludingly, evidence about the contextual interference effect in children with neuromotor disorders is still limited and unclear, findings from adult populations cannot be transferred to the paediatric population. With this knowledge, the obvious next step seemed to be to tackle the randomised controlled trial to gain information about this topic. As, in the beginning of this dissertation, we were looking for a way to perform the learning of motor tasks in a standardised way and to measure performance directly during the learning process, we decided to use a robotic device for the study. Also, financing of my PhD project was obtained for including robotics. At that time, the ChARMin robot was the youngest member of the robotic devices group and was described as being suitable for more severely affected children [21].

Hence, the choice fell on ChARMin. One disadvantage was that the psychometric properties of the ChARMin assessments were not yet determined. In order to have trustworthy results, reliable, valid and responsive assessments are needed. This made it necessary to include a separate reliability study of ChARMin's assessments as a part of this doctoral thesis.

ROBOTIC ASSESSMENTS

Reliability of the ChARMin assessments

During the second part of this PhD thesis (chapter 3), we thus evaluated the reliability of the six assessments (active and passive Range of Motion, Strength, Resistance to Passive Movement, Quality of Movement, Circle, and Workspace) provided by the ChARMin exoskeleton. We simultaneously collected data for the evaluation of the validity and the responsiveness of these assessments, yet, the number of participants and therewith the data for these two analyses are still limited. For the reliability study, we included 30 children with congenital (n=15) or acquired brain lesions (n=14), or both (n=1) and affected upper limb functions, representing the population undergoing upper limb neurorehabilitation at the Swiss Children's Rehab.

Both, the analyses of relative and absolute reliability showed a wide range of Intraclass Correlation Coefficients (ICC) and Smallest Real Differences, respectively, between the assessments and the parameters. This mainly is what we expected. Combined with the experiences we made, we found explanatory approaches for these differences.

Interestingly, while only single parameters of some assessments showed systematic errors between the measurements, the majority of the parameters of Quality of Movement indicated that the children performed the assessment more fluent, direct and fast during the second appointment. Yet, after correction for multiple comparisons, only the parameter 'mean time' (i.e. mean time needed to reach the target) shows a significant change. There seems to be some training effect, which could be responsible for the systematic error between the first and the second measurement. It would be interesting to assess the number of repetitions required until stable values are reached. With according familiarisation trials within the test protocol, a higher agreement between the first and the second measurement could be achieved.

The Range of Motion assessments proved not to be suitable to measure end of range movements as the motion range provided by ChARMin is limited due to safety reasons. The Resistance to Passive Movement assessment showed very low reliability and was therefore not recommended for the use in research projects and clinical practice. The Strength assessment showed promising aspects and we proposed some adaptations to further improve the reliability of the measurements.

For some more severely affected children, Quality of Movement and Circle were too difficult to perform. ChARMin was developed with support modes that enable the training specifically for more severely affected children [21]. Certainly, increasing the support against gravity would facilitate to reach the upper targets of the Quality of Movement assessment (eight targets are located radially around a centre point, reaching upper targets requires lifting the arm up against gravity, reaching the lower targets requires moving the arm downwards). For more severely affected children, the level of support could be adjusted to their physical abilities [21]. Yet, our clinical experience has shown that when increasing the level of support against gravity, it is difficult for the more severely affected children to reach the lower targets. To do so, the child has to move the arm downwards and against the support against gravity. For some participants, this was too much resistance to overcome.

We generally recommended the Quality of Movement, Circle and Workspace assessments for the use in clinical practice with children practicing with the ChARMin robot during therapy. Yet, it is important to be aware that reliability values obtained in a research setting (controlled environment, the same rater, the same daytime, etc.) is not exactly the same as in a clinical setting. There, all the mentioned factors might vary and influence the reliability of any measurement.

Implementation of robotic assessments in the clinical setting

Despite the increasing application of new technologies in paediatric neurorehabilitation [22], it has been stated that they are not well established in clinical practice [23]. In my opinion, new technologies such as robotic devices can and should be applied in paediatric neurorehabilitation if individual rehabilitation goals can be addressed with a robotic device (e.g. the goal is to improve precision of reaching, practicing ChARMin's Treasure Hunt exergame would be an option). Furthermore and more importantly, the child needs to like to practice with a robotic device. If this prerequisite is not given, using a robotic device is useless. The same applies when planning to use a robotic device as assessment tool. Additionally, conducting robotic assessments might be useful if the child is training with this same device. The set-up and adaptations of the device, cuffs, etc. is the same for the assessments and the training. Hence, it is worth to perform the rather time-consuming procedure when ChARMin is used for both, training and assessment, but it might not be when used only for an assessment. Furthermore, as seen in the reliability analysis of ChARMin's QoM assessment, a familiarisation phase is recommended for reliable measurements. Hence, before using data from the QoM assessment to evaluate a change due to an intervention (i.e. measurements before and after the intervention), the children should be familiar with the assessment. This would decrease the systematic error due to improvement with practice.

Despite the claim that the training with robotic devices in combination with exergames is task-specific [24], it has to be considered that a motor improvement obtained with

robotic training and measured with a robotic assessment not necessarily improves activities in daily life. According to the World Health Organisation's International Classification of Functioning, Disability and Health (ICF), activities in laboratory situations reflect the capacity (a person's ability to do something in a standard environment) while activities in a real life situation reflects the performance (what they actually do in their usual environment) [25]. To date, robotic training and assessments are mostly locally bound, reflecting rather a laboratory situation and not real life while conventional therapy can easier be applied in a setting which is closer to the normal life setting. Hence, robotic devices cover the capacity while conventional therapy has the opportunity to address the performance. Ideally, applying robotic devices goes hand-in-hand with conventional therapy. In my opinion, any kind of therapeutic approach, method, technique, and any device are therapeutic tools serving to attain the individual goals. The therapeutic reasoning should be the way being followed to decide about which tool should be applied and when.

While it might be a larger step to include robotic devices in clinical practice, for research, they provide a number of positive aspects: the opportunity to provide the same intervention several times (same intensity, time, numbers of repetitions, etc.) and to measure parametric data also during an intervention, are probably the most important ones. According to the experiences we gained during the reliability study, for our purposes, ChARMin seemed to be well suited. Next to assessments measuring aspects of motor learning (e.g. QoM, Circle), it provided exergames that should keep motivation high and that could be applied as motor learning tasks. Hence, we started with the next step, the evaluation of the effectiveness of blocked versus random order of motor tasks.

THE FEASIBILITY AND PILOT STUDY PROCESS

To assess the effectiveness of an intervention, usually a parallel or cross-over design randomised controlled trial comparing the intervention of interest with another intervention is performed. The randomised controlled trial design also is known as the "gold standard" for effectiveness research [26]. Yet, such a randomised controlled trial design comes with quite some expenditure and potential pitfalls, especially when being performed at a stage where aspects of the intervention itself are still at a developmental stage and the optimal outcome measure is not yet clear. Before spending resources of different origins only to receive unsatisfactory results, it is highly recommended to perform a feasibility or a pilot study first [27, 28]. However, research training does not or only rarely address the designing and conducting of pilot studies [33, 34]. Even the use of the terms "feasibility" and "pilot" is confusing and discussed controversially. Some experts use these two terms interchangeably [27], others distinguish between feasibility- and pilot studies [29]. It has been proposed that compared to a feasibility study, a pilot study has a stricter methodology

(including some justification for a sample size), an intention for future work, and it has been described as a smaller version of a future study (including randomisation procedure and a control group) which focuses mainly on the study processes [30]. Considering the abovementioned definitions and to follow a clear terminology in the feasibility/pilot process of this PhD project, the first experiment (chapter 4) was defined as "feasibility study", the second and more elaborate experiment (chapters 5 and 6) as "pilot study". It should be mentioned that here, the term "feasibility" is also used in the pilot study, as we have defined specific criteria with this pilot study to test the feasibility of a future main trial.

Feasibility study

We performed first measurements, mainly to explore the procedure of a potential study design. We included three participants and continuously analysed the findings. While during these first feasibility measurements, the planning of the study appointments additionally to the rehabilitation program was feasible, and participants were cognitively and physically able to play the exergames, the ChARMin assessments were a challenge for the participants due to physical and motivational limitations. Furthermore, technical issues with the ChARMin led repeatedly to missing data and negatively influenced the compliance of the participants. An equal number of repetitions of the exergames for each participant was not achieved.

With this feasibility trial, we gained new information but it also became obvious that we would need to follow a more guided way in order to approach a randomised controlled trial in the end. We would need a more elaborate study protocol including a sound reasoning for several choices and decisions: appropriate primary outcome measure and time point (i.e. acquisition, immediate or delayed transfer or retention), the interventions, numbers of sessions and repetitions, and the randomisation procedure. Retrospectively, already at this point, the inclusion of the participants would have been helpful: what do they consider important in terms of the primary outcome? How many repetitions of an exergame would be manageable for them, considering their compliance, motivation, interest, and/or fatigue?

Pilot study

While publishing pilot studies is considered important for research, no matter whether the outcome is positive or negative, [29] the publication of protocols of pilot studies is not very common. Publishing study protocols in general helps the readers to understand the study procedure better, enables them to implement the intervention and has become more common in recent years [31]. Particularly for a pilot study, it would be important to understand what has been done, as other researchers could learn from decisions and challenges, thus avoiding unnecessary negative duplication of work and making it possible to build on work already done.

We planned the pilot study (chapter 5) by elaborating a protocol. We aimed to address all the criteria needed to be fulfilled in order to conduct a future randomised controlled trial according to this protocol. We discussed unclear aspects in the study team (e.g. would the exergames be variable enough to generate a contextual interference effect?) also based on the evidence from literature. We came up with solutions for the open questions (e.g. practicing reaching movements with two spatially different variations of the same exergame). Furthermore, we defined specific criteria to evaluate the feasibility of a future main randomised controlled trial [27]. Motor learning outcomes were split in one primary and several secondary outcomes. We conducted the pilot study according to the a priori established study protocol. After performing one year of recruitment and measurements, the evaluation of the criteria showed that a future randomised controlled trial with the same parallel group design was not feasible. Despite this rather sobering conclusion, according to Thabane et al. this should not be considered a fail but rather a success since spending resources on a large study that will not be successful was avoided [27]. However, we have to develop new ways to evaluate contextual interference in children with brain lesions. The first step is to reconsider the choices made for the actual study.

In- and exclusion criteria

We defined the in-and exclusion criteria taking into account the heterogeneity of the population of children undergoing upper limb neurorehabilitation at the Swiss Children's Rehab. Additionally, the participants should have been able to perform the exergames and the assessments. The choice of the in- and exclusion criteria is critical and comes with some difficulties, as it was our experience. We chose the Manual Ability Classification System (MACS) level [32] as inclusion criterion for sufficient upper limb function. Classification systems as the MACS and the GMFCS have been recommended to serve as inclusion criteria since they are easily applied and reflect the functional performance on the activity level [33]. In several studies, the MACS was used as inclusion- or exclusion criterion [34–36]. Yet, for our study, we experienced that the MACS was not specific enough, as two participants who were not able to perform the exergames were included. It has been stated that the chosen eligibility criteria of physical activity are often too restrictive or not generalisable to the real world practice [28]. In our case it turned out that at least the MACS did not cover the conditions required for practicing the exergames well enough. While the MACS actually reflects the child's typical performance when using his or her hands (bimanual and collaborative) to handle objects in daily activities [32], during the ChARMin exergames, the child's capacity is required. The participants are asked to reach predefined targets as directly and quickly as possible. This may require extraordinary effort, range of motion or movement speed. To cover the ICF body function domain, we would need to implement a criterion of body function [25]. The

two participants dropping out due to physical inability to perform the exergames would have been detected if active range of motion of the upper limb joints and the strength of upper limb muscles would have been used as such a criterion.

We had to abort sessions and had drop-outs due to lack of compliance and motivation. There are many reasons, such as age, personality, cognition, interest for the exergames, the will to improve motor function, for being compliant or not. With the exception of age, it might thus be difficult or even impossible to evaluate the motivation or compliance in the sense of in- or exclusion criteria for a future procedure. An option would be to tackle the problem from the other end: if motivation and compliance are not given for long and challenging practice sessions, sessions could be shortened and in return the number of sessions could be increased to attain the required number of repetitions. This could, however, interfere with the rehabilitation schedule (length of stay), but it is an issue that must be considered for future research. Another approach would be to consult the children who would be the participants of this study: Asking them about what they would need to be compliant over a longer practice period could result in new ideas: what kind of exergame would motivate them? What would they need in terms of breaks between the trials? How should breaks be designed to provide a positive impact in the motivation, to reduce fatigue and increase compliance to go on with practice?

Study sample

We aimed to generate a study sample reflecting the heterogeneity of the population of children undergoing neurorehabilitation. Therefore, we included children with both congenital and acquired brain injuries in our pilot study. However, there are good reasons against this approach, and it is legitimate to discuss whether it makes sense. Generally speaking, a brain lesion leads to a damage of the brain and therewith its function. This leads to an impairment of motor, sensory and/or cognitive functions, whereby the affected functions and extent depends on the location and severity of the lesion. While “congenital” is defined as existing at or dating from birth [37], “acquired” means developed after birth [38]. Cerebral palsy, for example, is described as: “… a group of disorders of the development of movement and posture, causing activity limitation, that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain.” p. 572 [39]. The main difference between children with congenital and acquired brain lesions is the stage of brain development at the time of the incident. On the one hand, depending on the level of severity, children with congenital brain lesions have limited movement experience prior to the brain lesion that happened very early in their life. Most of the motor development takes place with a damaged central nervous system. Children with acquired brain lesion, on the other hand, had a normal development of the brain and motor functions and normal movement experience prior to the time point at which the brain lesion happened. However, depending on the time point of the lesion, the

level of movement experience varies greatly in the group of children with acquired brain lesions.

When considering paediatric neurorehabilitation, the question is, whether these two groups need to be treated differently. In this context, the distinction between habilitation and rehabilitation must be mentioned. Habilitation refers to attaining, retaining or to improve skills needed for daily life [40]. In the paediatric field, habilitation often is related to the development of motor skills that have not yet been accomplished [40]. Rehabilitation, as the prefix "re-" indicates, refers to re-attainment of skills, which have been lost or affected due to an injury or illness of the brain [40]. Hence, "habilitation" is mostly used in the context with congenital brain lesions and "rehabilitation" with acquired brain lesions [41]. But what does that mean in terms of motor learning? Are attainment and re-attainment of skills managed differently? Apparently, motor learning is required for both, habilitation [42] and rehabilitation [43]. But do both groups respond equally to the different motor learning principles such as in our case the practice order? Both groups seem to have intact implicit learning [44] which would also include procedural learning and therewith motor learning (see Figure 7.1).

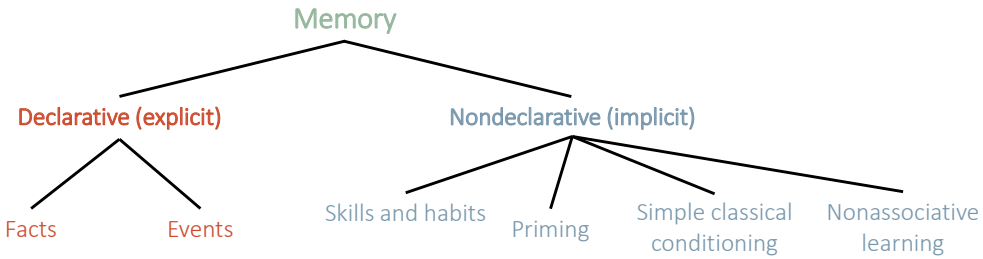


Fig. 7.1. Taxonomy of the human long-term memory. The memory of skills and habits is included in the implicit, nondeclarative memory. Redrawn from Squire & Zola-Morgan, 1990 [45].

The memory function does not seem to differ between the groups. However, the affected area of the brain is most certainly varying in most of the children, irrespective the aetiology (congenital or acquired) of the lesion. The cerebellum, parietal cortex, premotor cortex, motor cortex and the basal ganglia are the brain regions supposedly involved in motor learning and/or control [46]. Damage of one or several of these areas might affect motor learning. Differing results have been found about the impairment of motor learning in adult persons with hemiparesis after stroke [43]. As the aetiology and the affected brain area are only two of many differences between the individual patients after stroke, this is not surprising. It has been stated that finding a learning deficit in all patients after stroke might be unrealistic due to the heterogeneity of this population [43]. Given the similarly large heterogeneity in both, children with congenital and acquired brain lesions, this certainly also applies to these populations.

In the population of children with acquired brain lesions, the time that has passed since the event is a considerable factor too. In the first acute period after a traumatic brain injury or a stroke, the biological and spontaneous neurological recovery takes place. Functional improvements occur that do not necessarily have something to do with motor learning. A child with an acquired brain lesion in the acute phase therefore significantly differs compared to a child in a chronic, more stable phase after an acquired brain lesion. In addition, the psychological aspect of both, congenital and acquired brain lesions is not to be underestimated. A child after an acquired brain injury needs to get used to a completely new situation. The loss of the life as it was known before, needs to be processed. While a child with a congenital brain lesion has dealt with his or her body and its functions it has from the beginning of life, they see themselves and their difficulties in daily life in comparison to other, typically developing children. Paediatric traumatic brain injuries are associated with depression [47] and other psychiatric disorders [48], cerebral palsy with chronic pain, fatigue, and depressive symptoms [49]. Knowledge about the former life or the comparison to typically developing children can be a motivational factor, helping to stick to rehabilitation programmes to improve the situation. Children with acquired brain injuries might have a much clearer idea about what they want to re-attain, whereas children with congenital brain lesions have lived their whole lives with the opportunities they have. My clinical experience showed me that they are quite content with their situation: why should I work on my walking, I am very fast with my wheelchair. Yet, it is also imaginable that a comparison with the former life or the situation of typically developing children could have a depressing effect, influencing the rehabilitation progress negatively. Despite motivation being listed as a body function in the ICF [25], it also depends on the child's personality, the environment, support of family and friends and the given opportunities. The so called contextual factors (personal and environmental factors) are an important part of the ICF model, interacting with, in the case of motivation, a body function and therewith, the health condition of a person [25]. Hence, motivation for rehabilitation clearly is an individual matter and needs to be looked at as such.

Concluding, children with congenital and acquired brain injuries differ in some aspects. Yet, active therapies based on motor learning could be the best choice for both groups even if they might respond differently to the same treatment. It might just be irrelevant to group children with neuromotor disorders based on their diagnoses, expecting them to respond to treatments similarly. Once more, it seems to be recommended to look at this very heterogeneous population from a more individual point of view. This might apply to both, treatment and research.

Parametrisation of criteria and generalisability of results

Retrospectively, we should have parametrised the feasibility criteria more specifically. For example, we could have used a standardised measure to assess

the participants' fatigue and motivation. This would have provided more valid data for these parameters. Additionally, we could have chosen clear criteria for success [28]. For example, we could have predefined an acceptable recruitment rate or the amount of time we would accept for data processing. The conclusion about why we rated a criterion as not feasible would have been more comprehensible. Defining such criteria would enable unbiased decisions about the feasibility of a future main trial [27]. This was indeed a difficulty when analysing the data obtained during the pilot study. Such cut-off levels or criteria for success seem to be neglected in many pilot studies: Five of six analysed examples of pilot studies involving physical activity did not specify criteria for success [28]. Assumingly, the difficulty is to define such criteria as it requires a sound background of known data, which are not always available and results in a strict cut-off. In our study, the feasibility criteria were also interdependent, for example, a high adherence would not be helpful if we could only recruit a few participants. What if half of the criteria would recommend "feasible" and the other half "not feasible"? Such questions should be considered beforehand.

The generalisability of the results of this pilot study are limited. This PhD project was performed in the restricted and small setting of our rehabilitation centre, which makes comparisons and generalisability to other study results difficult. Outcomes of criteria such as the feasibility of the planning and the recruitment processes and the handling of the ChARMin robot for example are limited to our centre completely. In another setting, the randomisation procedure might be difficult to manage, for example due to lack of personnel and again in another setting, it might be feasible with only one person involved. Additionally, the recruitment rate is greatly influenced by the setting of our centre as the number of participants is restricted due to the number of available beds in the rehabilitation centre. In a larger clinic or in an outpatient setting, this criterion would possibly be feasible.

"**Can it work?**", "**Does it work?**", "**Will it work?**"

Another way to approach the different phases of a feasibility/pilot process suggests to ask three questions "Can it work?", "Does it work", and "Will it work?" and to define appropriate areas of interest [50]. The area of focus of the planned study and the research question point out appropriate study design options [50]. Areas of focus for feasibility and pilot studies are acceptability, demand, implementation, practicality, adaptation, integration, expansion, and limited efficacy [50]. The study proposing this approach rather focused on the feasibility of a new intervention within a certain setting [50]. Yet, this approach can also adjusted for the feasibility of conducting a study.

As the question "Will it work?" represents the real life situation or the effectiveness of an intervention [50], an area which has not been addressed in this PhD project, it will also not be discussed below. In our case, the first feasibility measurements can be allocated to the question "Can it work?" as during this phase the question

Table 7.1. Feasibility study design options for areas and research questions of interest

	Question of interest	**"Can it work?"**	**"Does it work?"**
	Phase of the feasibility/pilot process of this PhD project	Feasibility study (chapter 4)	Pilot study (chapters 5 and 6)
Areas of focus	Acceptability	Focus groups with target participants: How would this intervention fiot with daily-life activities?	Randomised controlled trial comparing the satisfaction of intervention and control group
	Implementation	Pre-post design Focus group as method of evaluation: Can the intervention be deployed in the clinical context?	Pre-post design Evaluation of small scale demonstration project Surveys and observations: Can the intervention be deployed in the clinical context? (Comparing practices and outcomes before and after intervention impelmentation
	Practicality	Small-scale demonstration study Key-informant interviews examining predicted cost, burden, and benefit because of appropriate intensity, frequency, duration of the intervention	Cost-effectiveness analysis Stakeholder interviews to determine how easily the intervention was used by the staff

Table adapted to the feasibility/pilot process of this thesis from Bowen et al., 2009 [50].

whether there is some evidence that the intervention might work is addressed [50]. The pilot study (chapters 5 and 6) is more relating to "Does it work?" as a potential efficacy of an intervention is more in the foreground [50]. Three areas of focus can be assigned to the feasibility/pilot process of this PhD projects: acceptability (reaction of participants and persons involved in the implementation of the programme to the intervention), implementation (extent, likelihood and manner in which an the intervention can be implemented as planned), and practicality (extent to which the

intervention can be delivered with constraint resources, time, and/or commitment) are areas of focus [50]. The recommended study design options for the appropriate combinations of the areas of focus and the questions relevant to the two phases of our feasibility/pilot process are shown in Table 7.1.

Our feasibility study was a small-scaled study, which is recommended to answer the question "Can it work?" for practicality (Table 7.1). It also has been proposed that focus groups (acceptability and implementation) and key informant interviews (practicality) would be appropriate methods of evaluation in his phase [50]. In recent years, participatory research approaches including the views, knowledge, experience and action of persons to whom the research is focused on [51].

However, children are still mainly involved in research as participants and not as partners [52, 53]. We did not include the participants' voices about their impression but our perception thereof by using notes we had taken during the measurements about the behaviour of the subjects as a basis for our evaluation. It certainly would be highly educational to involve the children as the target population of such interventions and or their parents already from the beginning of a research process.

IMPLICATIONS FOR FUTURE RESEARCH

Randomised controlled trial: the solution or the problem?

Initially, at the beginning of this PhD project, we planned to conduct a main randomised controlled trial evaluating the contextual interference effect. During the process, however, we altered this objective as methodological issues needed to be addressed before we could continue with an interventional trial focused on effectiveness. It has been recommended to place the main focus on the feasibility rather than the effectiveness in a pilot study [27], which is what we did. An important and necessary step to take in the whole process. Yet, it meant to completely change the focus from motor learning to methodological characteristics of the trial.

While we were very critical about single aspects of the study design (i.e. which primary outcome, what kind of intervention, numbers of repetitions), we did not question the choice of a randomised controlled trial design per se. Randomised controlled trials have been determined to be the most stringent designs to assess whether there is a cause-effect relationship between the outcome and the intervention [54]. They are claimed to be the gold standard for effectiveness studies [26]. However, in our small and heterogeneous setting of paediatric neurorehabilitation, it is very difficult to recruit the required number of participants to enable an according power. We experienced this during the recruitment phase of our pilot study. Also, randomised controlled trials mostly only include a small number of all the different characteristics of children, limiting the sample size and therewith the generalisability even more [55].

Yet, is important to look at randomised controlled trial designs with a more differential look. According to the science philosopher Nancy Cartwright, methods that are less rigorous on internal validity might be justified much better in terms of external validity and she states that trade-offs between internal and external validity need to be considered [56]. Also, she suggests that there is no reason to choose a method that is stringent on one end (internal validity) and rather unclear when it comes to the generalisability over a method that is less rigorous but reasoned throughout the whole study [56]. Explanatory trials have a high internal validity. With a stringent design, group differences can be related to the intervention [57]. Pragmatic trials have a high external validity, yielding results that can be generalised to a setting outside of research. In other words: „There is no gold standard." p.11 [56]. The research question thus decides upon the study design: If internal validity is needed to address the research question, an explanatory trial is the choice, if external validity with a high generalisability is required, a pragmatic trial is the research design of choice.

Hence, with the knowledge we have now, designs that are maybe less strict when it comes to internal validity but enable to transfer findings into the real world setting need to be considered to evaluate contextual interference in children with brain lesions. Additionally, there are many known and unknown factors influencing contextual interference and no universal practice schedule can be accounted as beneficial in all contexts and types of persons [6]. A more individual approach is thus required.

Alternative research designs for paediatric neurorehabilitation

The studies of this PhD project led to many insights but until now, it is unclear, whether children with brain lesions would respond to the contextual interference effect similar as healthy adults. As it became obvious with the findings of the feasibility and pilot studies, alternative study designs to a randomised controlled trial need to be considered. Recently, single-subject designs became more important also in the field of paediatric rehabilitation. The advantage of a single-subject design is the variability of responses, which is considered between and within the participants [58]. This can be very helpful because especially in paediatric rehabilitation, the study samples are often very heterogeneous [58]. The main question that can be answered with a single subject design is the following: "Are changes in the independent variable (the characteristic or intervention of choice) associated with measurable change in the outcome of interest (dependent variable)?" p.576 [58]. Basically, this is the question we would like to have answered: Is the practice order associated with changes in the motor learning outcome? With two practice order groups an ABACA design with A representing the baseline (and possibly transfer measures), B is the blocked order practice and C the random order practice, would be a possibility. Important would be that the order of the practice groups would be randomised between the participants and a washout phase would be needed between B and C in order to prevent from

a carry-over effect [59]. Yet, as a long-lasting retention effect is desired, this could pose a conflict, needed to be solved. Next to this, also other potential difficulties need to be considered. The average length of stay of participants undergoing neurorehabilitation is approximately 40 days. This limits the duration of the study, unless, of course, we would choose treatment and assessment options that can also be conducted at the participants' homes. This would enable the inclusion of outpatient participants as well. Furthermore, the number of assessment time points would increase as this is an important feature of a single-subject design. In order to meet evidence standards, a reversal/withdrawal study needs to have at least five data points in each phase [60].

A recently published review about single-subject design studies with children with cerebral palsy stated that the studies included in their review, generally had a high overall risk of bias [55]. This is disappointing because according to the authors, this type of study has potential in the context of evidence-based medicine in children with cerebral palsy [55]. They strongly recommend to closely follow the guidelines and to build a highly valid and clinically relevant base of evidence [55]

Eventually, the translation of research and the application and implementation of study findings in the clinical practice is the first step to the most important goal of research, namely to generally improve the situation of the children with neuromotor disorders in their daily life. Most importantly, what should change in research in general, is the inclusion of the persons concerned from the beginning. With qualitative research approaches, relevant research questions could be elicited. They should come from the children undergoing neurorehabilitation, their parents and siblings, caregivers or therapists. Involving these groups of people from the beginning and throughout the whole research process, makes research relevant, meaningful, and facilitates the knowledge transfer. For example, it is of great interest to evaluate what participants need in order to be motivated and compliant for motor learning sessions. What is of interest for them? What is their goal, a more fluent movement or rather the ability to grasp an object without breaking it? Where are their mental and physical limits? How long can they concentrate on something? Addressing the children themselves, their parents, teachers, therapists and caregivers with these kinds of questions first with a qualitative method followed by a quantitative study evaluating these aspects, would elicit the whole situation more deeply. Combining both, qualitative and quantitative data, provides better understanding of a research problem than either part for itself [61]. Depending on the research questions and the rationale different mixed methods designs (e.g. triangulation-, exploratory-, explanatory-, or embedded designs [62]) are usually applied. This approach might reflect paediatric neurorehabilitation more realistically and in the end it also should be a pragmatic way to go.

CONCLUSION

This PhD thesis evaluated the current state of the evidence about contextual interference in children and revealed a lack of methodological sound studies. The reliability of the ChARMin assessments was assessed and recommendations about the use of these assessments were given. The most important conclusion, however, was elicited by performing the feasibility and pilot process of evaluating the effectiveness of blocked versus random practice order of two exergames in children with brain lesions. The paediatric neurorehabilitation needs more alternative research approaches. Good-quality randomised controlled trials focussing on effectiveness with optimal preparation of the feasibility of all aspects of the study are very difficult to achieve in this field: the population is heterogeneous, limiting this heterogeneity would result in a loss of important features of this population and even lower numbers of participants. In our setting, it is not possible to recruit a study sample large enough to generate an appropriate power. Hence, more individual approaches such as single-subject designs need to be considered in order to provide meaningful research. To cover the relevance of our research, the inclusion of the persons of interest, children, their parents and siblings, friends and caregivers is needed. The focus of research approaches needs to change, a mixed-method design might provide a deeper understanding of the issue about how to address the individuality of the participants in studies. In the end, what we do should be for the benefit of the children and the least we could do is to involve them more in research to make the outcome more meaningful to them.

REFERENCES

1. Duff S V, Gordon AM. Learning of grasp control in children with hemiplegic cerebral palsy. Dev Med Child Neurol. 2003;45:746–57.
2. Van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine (Phila Pa 1976). 2003;28:1290–9.
3. Figueiredo TP, Barreiros JMP. Contextual interference research: An overview. Mot Humana. 1993;9:61–71.
4. Magill RA, Hall KG. A review of the contextual interference effect in motor skill acquisition. Hum Mov Sci. 1990;9:241–89.
5. Brady F. A theoretical and empirical review of the contextual interference effect and the learning of motor skills. Quest. 1998;50:266–93.
6. Merbah S, Meulemans T. Learning a motor skill: Effects of blocked versus random practice- A review. Psychol Belg. 2011;51:15–48.
7. Cauraugh JH, Kim SB. Stroke motor recovery: Active neuromuscular stimulation and repetitive practice schedules. J Neurol Neurosurg Psychiatry. 2003;74:1562–6.
8. Hanlon RE. Motor learning following unilateral stroke. Arch Phys Med Rehabil. 1996;77:811–5.
9. Schweighofer N, Lee J, Goh H, Choi Y, Kim SS, Stewart JC, et al. Mechanisms of the contextual interference effect in individuals poststroke. J Neurophysiol. 2011;106:2632–41.
10. Jo E-J, Noh D-H, Kam K-Y. Effects of contextual interference on feeding training in patients with stroke. Hum Mov Sci. 2020;69:1–9.
11. Ceci A, Giannuzzi V, Bonifazi D, Felisi M, Bonifazi F, Ruggieri L. Clinical trials in paediatrics - regulatory and methodological aspects. In: Drug Discovery and Development- From Molecules to Medicine. IntechOpen; 2015. p. 271–97.
12. Joseph PD, Craig JC, Caldwell PHY. Clinical trials in children. Br J Clin Pharmacol. 2013;79:357–69.
13. [no author listed]. Time to be serious about children's health care. Lancet. 2001;358:431.
14. Klassen TP, Hartling L, Craig JC, Offringa M. Children are not just small adults: The urgent need for high-quality trial evidence in children. PLoS Med. 2008;5:1180–2.
15. Leversen JSR, Haga M, Sigmundsson H. From children to adults: Motor

performance across the life-span. PLoS One. 2012;7.

16. Cheyne D, Jobst C, Tesan G, Crain S, Johnson B. Movement-related neuromagnetic fields in preschool age children. Hum Brain Mapp. 2014;35:4858–75.

17. Largo RH, Fischer JE, Rousson V. Neuromotor development from kindergarten age to adolescence: Developmental course and variability. Swiss Med Wkly. 2003;133:193–9.

18. Musselman KE, Roemmich RT, Garrett B, Bastian AJ. Motor learning in childhood reveals distinct mechanisms for memory retention and re-learning. Learn Mem. 2016;23:229–37.

19. Lejeune C, Catale C, Schmitz X, Quertemont E, Meulemans T. Age-related differences in perceptuomotor procedural learning in children. J Exp Child Psychol. 2013;116:157–168.

20. Tillman CM, Thorell LB, Brocki KC, Bohlin G. Motor response inhibition and execution in the stop-signal task: Development and relation to ADHD behaviors. Child Neuropsychol. 2007;14:42–59.

21. Keller U, Van Hedel HJA, Klamroth-Marganska V, Riener R. ChARMin: The first actuated exoskeleton robot for pediatric arm rehabilitation. IEEE/ASME Trans Mechatronics. 2016;21:2201–13.

22. Falzarano V, Marini F, Morasso P, Zenzeri J. Devices and protocols for upper limb robot-assisted rehabilitation of children with neuromotor disorders. Appl Sci. 2019;9.

23. Iandolo R, Marini F, Semprini M, Laffranchi M, Mugnosso M, Cherif A, et al. Perspectives and challenges in robotic neurorehabilitation. Appl Sci. 2019;9.

24. Klamroth-Marganska V, Blanco J, Campen K, Curt A, Dietz V, Ettlin T, et al. Three-dimensional, task- specific robot therapy of the arm after stroke: a multicentre, parallel-group randomised trial. Lancet Neurol. 2014;13:159–66.

25. World Health Organization. International classification of functioning, disability and health - children & youth version. Geneva, CH: WHO Press, World Health Organization; 2007.

26. Hariton E, Locascio JJ. Randomised controlled trials—The gold standard for effectiveness research. An Int J Obstet Gynaecol. 2018;125.

27. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: The what, why and how. BMC Med Res Methodol. 2010;10.

28. El-Kotob R, Giangregorio LM. Pilot and feasibility studies in exercise, physical activity, or rehabilitation research. Pilot Feasibility Stud. 2018;4.

29. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol. 2010;10.

30. Whitehead AL, Sully BGO, Campbell MJ. Pilot and feasibility studies: Is there a difference from each other and from a randomised controlled trial? Contemp Clin Trials. 2014;38:130–3.

31. Thabane L, Lancaster G. A guide to the reporting of protocols of pilot and feasibility trials. BMC Pilot Feasibility Stud. 2019;8:5–7.

32. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Ohrvall A-M, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: Scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006;48:549–54.

33. Güçhan Z, Mutlu A. The effectiveness of taping on children with cerebral palsy: A systematic review. Dev Med Child Neurol. 2016;59:26–30.

34. Albiol-Pérez S, Gómez J-AG, Olmo E, Soler AM. A virtual fine rehabilitation system for children with cerebral palsy: Assesment of the usability of a low-cost system. In: Rocha Á, Correia AM, Adeli H, Reis LP, Costanzo S, editors. Recent advances in information systems and technologies, Volume 2. Cham: Springer International Publishing; 2017.

35. Gerber CN, Plebani A, Labruyère R. Translation, reliability, and clinical utility of the Melbourne Assessment 2. Disabil Rehabil. 2019;41:226–34.

36. Jackman M, Novak I, Lannin N. Effectiveness of functional hand splinting and the cognitive orientation to occupational performance (CO-OP) approach in children with cerebral palsy and brain injury: Two randomised controlled trial protocols. BMC Neurol. 2014;14.

37. Dictionary Merriam Webster. Congenital. 2020. https://www.merriam-webster.com/dictionary/congenital. Accessed 12 Mar 2020.

38. Dictionary Merriam Webster. Acquired. 2020. https://www.merriam-webster.com/dictionary/acquired. Accessed 12 Mar 2020.

39. Bax M, Goldstein M, Rosenbaum P, Leviton A, Paneth N, Dan B, et al. Proposed definition and classification of cerebral palsy, April 2005. Dev Med Child Neurol. 2005;47:571–6.

40. Neurological and physical abilitation centre. The difference between habilitation and rehalbilitation. 2020. https://www.napacenter.org/?s=habilitation. Accessed 12 Mar 2020.

41. Kerbl R, Sperl W, Strassburg HM, Pettoello-Mantovani M, Ehrich J. Overview of

habilitation and rehabilitation for children and adolescents in Europe. J Pediatr. 2016;172:233–5.

42. Krebs HI, Fasoli SE, Dipietro L, Fragala-Pinkham M, Hughes R, Stein J, et al. Motor learning characterizes habilitation of children with hemiplegic cerebral palsy. Neurorehabil Neural Repair. 2012;26:855–60.

43. Kitago T, Krakauer JW. Motor learning principles for neurorehabilitation. In: Barnes MP, Good DC, editors. Handbook of Clinical Neurology. Amsterdam: Elsevier; 2013. p. 93–103.

44. Yeates KO, Enrile BG. Implicit and explicit memory in children with congenital and acquired brain disorder. Neuropsychology. 2005;19:618–628.

45. Squire LR, Zola-Morgan S. The medial temporal lobe memory system. Science. 1991;253:1380–6.

46. Shadmehr R, Krakauer JW. A computanional neuroanatomy for motor control. Exp Brain Res. 2018;185:359–81.

47. Laliberté Durish C, Pereverseff RS, Yeates KO. Depression and depressive symptoms in pediatric traumatic brain injury: A scoping review. J Head Trauma Rehabil. 2018;33.

48. Bloom DR, Levin HS, Ewing-Cobbs L, Saunders AE, Song J, Fletcher JM, et al. Lifetime and novel psychiatric disorders after pediatric traumatic brain injury. J Am Acad Child Adolesc Psychiatry. 2001;40:572–9.

49. van der Slot WMA, Nieuwenhuijsen C, van den Berg-Emons RJG, Bergen MP, Hilberink SR, Stam HJ, et al. Chronic pain, fatigue, and depressive symptoms in adults with spastic bilateral cerebral palsy. Dev Med Child Neurol. 2012;54:836–42.

50. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med. 2009;36:452–7.

51. Larsson I, Staland-Nyman C, Svedberg P, Nygren JM, Carlsson I-M. Children and young people' s participation in developing interventions in health and well-being: A scoping review. BMC Heal Serv Resreach. 2018;18.

52. Daly W. "Adding their flavour to the mix": Involving children and young people in care in research design. Aust Soc Work. 2009;62:460–75.

53. Hill M. Children's voices on ways of having a voice Children's and young people's perspectives on methods used in research and consultation. Childhood. 2006;13:69–89.

54. Sibbald B, Roland M. Why are randomised controlled trials important? Br Med J. 1998;316:201.

55. Beckers LWME, Stal RA, Smeets RJEM, Onghena P, Bastiaenen CHG. Single-case design studies in children with cerebral palsy: A scoping review. Dev Neurorehabil. 2020;23:73–105.

56. Cartwright N. Are RCTs the gold standard? Biosocieties. 2007;2:11–20.

57. Spieth PM, Kubasch AS, Penzlin AI, Illigens BM-W, Barlinn K, Siepmann T. Randomized controlled trials-a matter of design. Neuropsychiatr Dis Treat. 2016;12:1341–9.

58. Romeiser-Logan L, Slaughter R, Hickman R. Single-subject research designs in pediatric rehabilitation: A valuable step towards knowledge translation. Dev Med Child Neurol. 2017;59:574–80.

59. Kravitz R, Naihua D, Sunita V, Jiang L, The DEcIDE Methods Center N-0f-1 Guidance Panel. Introduction to N-of-1 trials: Indications and barriers. In: RL K, Duan N, editors. Design and implementation of N-of-1 trials: a user's guide. Rockville: Agency for Healthcare Research and Quality; 2014. p. 1–11.

60. Kratochwill TR, Hitchcock J, Horner RH, Levin JR, Odom SL, Rindskopf D, et al. Single-case designs technical documentation. What Works Clearinghouse. 2010. https://files.eric.ed.gov/fulltext/ED510743.pdf.

61. Creswell JW. Educational research: Planning, conducting and evaluating quantitative and qualitative research. 4th editio. Boston: Pearson Education, Inc.; 2012.

62. Plano Clark VL, Huddleston-Casas CA, Churchill SL, O'Neil Green D, Garrett AL. Mixed methods approaches in family science research. J Fam Issues. 2008;29:1543–66.

ADDENDUM

SUMMARY
ZUSAMMENFASSUNG
IMPACT
ACKNOWLEDGEMENT
ABOUT THE AUTHOR
LIST OF PUBLICATIONS

SUMMARY

Motor learning principles form the basis of paediatric neurorehabilitation. To provide treatment with the best possible outcome for children with congenital or acquired brain lesions, considering different aspects influencing motor learning is important. One of the motor learning principles refers to the practice schedule. If several motor tasks are practiced within the same session, they may interfere with each other, leading to the so called contextual interference effect. The contextual interference effect is the main topic of this dissertation.

Chapter 1 introduces the reader to motor function and learning in general, the contextual interference effect, paediatric neurorehabilitation and specifically new technologies applied in paediatric neurorehabilitation.
Motor function is a key prerequisite for human social interaction. It can be divided into motor development, motor control, and motor learning. Motor development is a process of individual maturation. Motor control, which is described as the influence of neurophysiological factors on human movement is affected by the maturation of different systems on the physiological levels. Motor learning results in a relatively robust gain in skilled motor performance and occurs through practice. In this chapter, the rather complex construct "motor learning" is illuminated from different points of view. Different motor learning explanatory approaches are discussed as well as the neurophysiological background of motor learning. Since motor learning literally happens in the brain, it is not possible to observe the process per se from the outside. However, its products are observable as a change in performance. The question at what time point of the learning process it can be stated that motor learning has occurred is discussed in this chapter. Furthermore, the terms acquisition, retention and transfer are explained.

Chapter 1 also describes paediatric neurorehabilitation in Switzerland. The largest part of the inpatient population of the only paediatric rehabilitation centre in Switzerland, the Swiss Children's Rehab, University Children's Hospital Zurich, are children with cerebral palsy. Alongside conventional rehabilitation therapies, a specialty of the Swiss Children's Rehab is the application of new technologies. Chapter 1 presents an overview of robot-assisted upper limb training in paediatric neurorehabilitation. Detailed information about the ChARMin robot, an exoskeleton providing exergames and assessments to practice and measure upper limb function, is given. New technologies, specifically robotic devices combined with exergames, can provide a large number of repetitions during the same treatment session and are promising for motor learning and therewith neurorehabilitation. Yet, chapter 1 also describes challenges especially in paediatric neurorehabilitation: it becomes boring to repeat the same task over and over again, so, to improve motivation, a therapist can start to include other tasks in the same session. This is where the contextual interference

effect, comes into play. If several motor tasks are practiced during the same session, they can functionally interfere with each other. If the tasks are practiced in blocked order (e.g. AAA…, BBB…., CCC….) the interference is low compared to when practiced in random order (e.g. ABBCABCAC…), then the interference is high. The contextual interference effect mainly has been assessed in healthy adults and has an impact on motor learning. High contextual interference (random practice order) leads to a better motor performance during retention and transfer compared to low contextual interference (blocked practice order). During the acquisition, blocked practice order leads to better motor performance compared to random practice order.
The chapter ends with the personal motivation of the author to perform this doctoral thesis followed by the outline and aim of the thesis. The general aim was to investigate the contextual interference effect in children with motor impairments due to brain lesions.

Within **Chapter 2**, a systematic literature search and evaluation of evidence about contextual interference in typically developing children and children with congenital or acquired brain lesions is presented. The systematic review provides an overview of 25 papers, covering 27 experiments, assessing the contextual interference effect (blocked versus random order practice in typically developing children). Generally, the methodological quality of the included studies was low and the risk of bias high. Only one study examined a group of children with cerebral palsy together with a group of typically developing children. The study did not present the results from only the group of children with cerebral palsy. A best evidence synthesis of the study results of typically developing children was performed. It resulted in limited evidence for the benefit of blocked practice over random practice for three tasks concerning the acquisition. Analyses of the retention reported limited evidence for the random over blocked practice for two tasks. Analysis of the transfer results reported moderately consistent evidence for the benefit of random over blocked practice for two tasks.

Within this chapter, the influence of types of skills and variations, experience and age of the learner, and task difficulty on the contextual interference effect is discussed. Furthermore, the question whether the knowledge about the contextual interference effect obtained in related populations related (i.e. adults with brain lesions, typically developing children) can be transferred to children with brain lesions, is addressed. For example, compared to typically developing children, physical requirements are different in the other two groups. Due to brain damage, error detection and consequently motor learning can be impaired. Certainly, further research is needed about this topic to fully understand the impact of pathological changes on motor learning.

In **Chapter 3**, a study about the reliability of and lessons learnt from assessments provided by ChARMin, an exoskeleton to practice and measure upper limb function,

is presented. This study was performed because we planned to apply ChARMin as a device for children with brain lesions to practice several motor tasks in either blocked or random order.

In – and outpatients of the Swiss Children's Rehab, aged between five to 18 years, with congenital or acquired brain lesions affecting upper limb function attended two appointments. Both appointments included the same assessments in a random order. Seven ChARMin assessments were performed. The assessments measured the active range of motion (aROM), the passive range of motion (pROM), strength (Strength), resistance to passive movements (RPM), the quality of goal-directed movements (QoM), the tracking capability of the arm (Circle), and the maximally reachable workspace (Workspace). Relative reliability was determined by calculating Intraclass Correlation Coefficients (ICC). To determine the absolute reliability, the Smallest Real Difference as an absolute ratio of the grand mean (|SRD%|) was calculated.

The relative and absolute reliability of the assessments and their individual parameters varied widely. The highest ICC was obtained for the 'maximum distance to front' parameter of the Workspace assessment (ICC = 0.95, 95% confidence interval (95%CI) [-0.41, 0.36]).The lowest ICC was obtained for the 'resistance against shoulder external rotation' parameter of the RPM assessment (ICC =- 0.03, 95%CI [-0.41, 0.36]). The lowest |SRD%| was obtained for the pROM parameter 'shoulder extension' (5.9%), while the highest one was registered for the RPM parameter 'resistance against forearm pronation' (41'810.1%).

Within this chapter, results were discussed and compared with existing literature. Additionally to the reliability results, the application of these assessments revealed some important issues that were taken into account in our recommendations for the clinical use of the assessments. It was concluded that the application of aROM and pROM is not recommended since the range of motion is limited mechanically by the device due to safety regulations. RPM also is not recommended due to the consistently poor reliability results and uncertainties about what construct RPM is actually measuring. Strength shows potential for clinical use, yet, it is recommended to adapt the measuring procedure and data evaluation processes. We recommend using the QoM, Circle, and Workspace assessments in clinical practice. Most of their parameters showed acceptable reliability and there are no alternative conventional assessments evaluating similar parameters. Yet, caution is advisable, as more severely impaired children had difficulties in performing the QoM and Circle assessments.

In **Chapter 4**, the first feasibility experiments comparing blocked and random practice order of exergames provided by ChARMin are presented. It was a limited and straight forward approach with the main objective to gain some information about participants' acceptability, implementation of the intervention, and practicality

of a future study procedure. Children with congenital or acquired brain lesions undergoing neurorehabilitation were recruited to practice two different exergames, either in blocked or in random order on two occasions. The first appointment was planned during the second to last therapy session of the day, the second appointment during the first therapy session the next morning. Before and after both sessions and one hour after the first session, assessment blocks (Box and Blocks Test, QoM and Circle assessments provided by ChARMin) were performed. We categorised notes taken throughout the study procedure into 'acceptability of the participants', 'implementation of the intervention in the clinical context', and 'practicality of the study procedure'. The main issues of the procedure were the technical problems we experienced with the ChARMin device (e.g. errors leading to interruption of the use of the device). Number and severity of these problems seemed to be closely related to the compliance of the participants. These results and additional aspects of the implementation of the study and its single components are discussed. At the end of the chapter, we concluded that extensive adaptations will be required in order to continue in a scientific more sound study built on this one. Thus, we first discussed adaptations and troubleshooting together with the ChARMin engineer. Based on the results of this experiment, we also developed a detailed study protocol for a pilot study about the contextual interference effect in children with either congenital or acquired brain lesions by applying either blocked or random practice of robotic exergames.

In **Chapter 5**, the abovementioned detailed study protocol of a pilot study investigating blocked versus random practice order of an upper limb robotic exergame, is described. This chapter points out that knowledge about the contextual interference effect in children with congenital and acquired brain lesion is of interest to potentially improve neurorehabilitation outcome in these children. In this chapter, we present the protocol of a pilot study by evaluating ten feasibility questions:

1. Are the inclusion criteria specific enough to result in a sample of participants, which is suitable for this pilot study (e.g. are they able to play the exergame)?
2. Is the recruitment rate feasible?
3. Is the scheduling procedure feasible?
4. Is the randomisation procedure feasible?
5. Is the whole procedure feasible for the participants?
6. Is the handling of the robot feasible?
7. Is the handling of the large amount of data feasible?
8. Are the outcome measures responsive and sensitive enough within this setting?

9. Is there a confounding influence of parallel therapies within the rehabilitation setting?
10. Is it feasible to conduct the main trial with respect to the needed sample size calculated from the data obtained for the primary motor learning outcome?

We chose an exergame that can be played with upper limb movements in the frontal or horizontal plane. The player is required to perform reaching movements in different directions to reach eight goals per trial appearing radially around a centre point. By practicing these variations (i.e. moving in different planes) of the exergame, we aimed to induce different levels of contextual interference (i.e. low level of contextual interference through blocked practice order, high level of contextual interference through random practice order). A three-week study procedure was planned: During week 1 (control week) two assessment blocks were planned. Week 2 (practice week) consisted of three practice sessions during which participants would perform 30 trials of each version of the exergame. Immediately before the first practice session, immediately after, one hour after, one day after and one week after the last practice session, assessments would be performed. Hence, week 3 served as follow-up week. Chapter 5 closes with a discussion about the wider context in which the pilot study can be set: it is a first step towards getting more information about the contextual interference effect in children with congenital or acquired brain lesions.

Within **Chapter 6**, results of the pilot study as described in chapter 5 are presented. We recruited participants among the inpatients of the Swiss Children's Rehab during one year. The order of the ten feasibility questions was adjusted according to the importance of the questions (the first question is the most important). The analysis of the data about the ten questions resulted in the following answers:

1. The in- and exclusion criteria were not specific enough as two participants were physically not able to play the exergames.
2. Recruitment rate: We informed 18 eligible children about the study, 12 of them agreed to participate and were enrolled. There were several drop-outs at different stages of the study for several reasons. In the end, four complete data sets and two incomplete data sets (missing follow-up measurement, data of week 1 excluded) were available. This rate is too low and not feasible.
3. A sample size calculation based on collected data of the this study reported that a future main trial would require 15`355 participants and is therefore not feasible.
4. For 50% of the participants, the procedure was not feasible, as they were not able to perform 30 trials of each exergame during each session.

5. The scheduling procedure of the study related sessions was difficult, mainly because the study appointments needed to be planned between the regular therapy appointments of the individual children. Mandatory rehabilitation appointments planned on short notice or illness of a participant were additional obstacles.
6. The randomisation procedure worked well, and the randomisation by minimisation programme could be applied without problems.
7. The handling of the robot proved to be better compared to the first feasibility experiments (chapter 4), yet, there still were technical issues that could have influenced the data.
8. The amount of data was feasible, yet a more advanced self-regulating way to analyse the ChARMin data would facilitate the procedure.
9. Outcome measures were only weakly responsive, yet, these results need to be treated cautiously as they were calculated on the basis of limited data availability.
10. Parallel therapies did not seem to influence the results. Yet, also here, the underlying amount of data is limited.

Chapter 6 discusses the feasibility results and compares them with existing literature. Concludingly, a main trial with the procedure of this pilot study is not feasible. Maybe an alternative study design might be more beneficial to gather information about the contextual interference effect in children with brain lesions.

In **Chapter 7**, the general discussion of this dissertation, the main results and critically discusses specific aspects of the studies included in this dissertation are discussed. The main focus lies on the feasibility and pilot study process. Depending on literature, the denominations “feasibility study” and “pilot study” are used interchangeably or sometimes are given different meanings: Compared to a feasibility study, a pilot study has a more stringent protocol and an intention for future work, namely a main trial.

It has also been suggested to use the questions “Can it work?” “Does it work?” “Will it work?” to plan and report a feasibility/pilot study process. The forerunner study described in chapter 4 (“Can it work?”) and the pilot study described in chapters 5 and 6 (“Does it work?”) are inserted in the framework of these questions. Chapter 7 finishes with a paragraph about implications for future research. Based on the experiences we made throughout this thesis, it is questionable, whether a randomised controlled effectiveness trial actually is suitable to assess the contextual interference effect in paediatric neurorehabilitation. This issue is discussed and suggestions for alternative study designs that consider more the individuality of the participants (e.g. single subject design studies, mixed methods design studies) are made. Additionally, it is emphasized that the important stake holders (namely

representatives of the population of children undergoing neurorehabilitation, their parents, siblings, friends and caregivers) should be included more to increase the relevance of the research.

ZUSAMMENFASSUNG

Prinzipien des motorischen Lernens bilden die Grundlage der pädiatrischen Neurorehabilitation. Um eine Behandlung mit dem bestmöglichen Ergebnis für Kinder mit angeborenen oder erworbenen Hirnläsionen zu gewährleisten, ist die Berücksichtigung verschiedener Aspekte, die das motorische Lernen beeinflussen, wichtig. Ein Prinzip des motorischen Lernens bezieht sich auf die Übungsreihenfolge. Wenn mehrere motorische Aufgaben in der gleichen Sitzung geübt werden, können sie sich gegenseitig beeinflussen, was zum so genannten kontextuellen Interferenzeffekt führt. Der kontextuelle Interferenzeffekt ist das Hauptthema dieser Dissertation.

Kapitel 1 führt den Leser in die Motorik und das Lernen im Allgemeinen, den kontextuellen Interferenzeffekt, die pädiatrische Neurorehabilitation und speziell in neue Technologien ein, die in der pädiatrischen Neurorehabilitation eingesetzt werden. Die motorische Funktion ist eine wichtige Voraussetzung für menschliche soziale Interaktion. Sie kann in motorische Entwicklung, motorische Kontrolle und motorisches Lernen unterteilt werden. Die motorische Entwicklung ist ein Prozess der individuellen Reifung. Die motorische Kontrolle, die als Einfluss neurophysiologischer Faktoren auf die menschliche Bewegung beschrieben wird, wird durch die Reifung verschiedener Systeme auf physiologischen Ebenen beeinflusst. Motorisches Lernen führt zu einem relativ robusten Zuwachs an geschickter motorischer Leistung und erfolgt durch Übung. In diesem Kapitel wird das recht komplexe Konstrukt “motorisches Lernen” aus verschiedenen Blickwinkeln beleuchtet. Verschiedene Erklärungsansätze des motorischen Lernens werden ebenso diskutiert wie dessen neurophysiologischer Hintergrund. Da motorisches Lernen buchstäblich im Gehirn stattfindet, ist es nicht möglich, den Prozess per se von aussen zu beobachten. Seine Produkte sind jedoch als Leistungsveränderung beobachtbar. Die Frage, zu welchem Zeitpunkt des Lernprozesses man feststellen kann, dass motorisches Lernen stattgefunden hat, wird in diesem Kapitel diskutiert. Weiterhin werden die Begriffe Acquisition (Erwerb), Retention (Behalten) und Transfer (Übertragung) von motorisch Gelerntem erläutert.

In Kapitel 1 wird auch die pädiatrische Neurorehabilitation in der Schweiz vorgestellt. Den grössten Teil der stationären Patienten des einzigen pädiatrischen Rehabilitationszentrums der Schweiz, der Kinder Reha Schweiz, Universitätskinderspital Zürich, sind Kinder mit Zerebralparese. Neben den konventionellen Rehabilitationstherapien ist die Anwendung neuer Technologien eine Spezialität der Kinder Reha Schweiz. Kapitel 1 gibt einen Überblick über das robotergestützte Training der oberen Extremitäten in der pädiatrischen Neurorehabilitation. Detaillierte Informationen über den ChARMin-Roboter, ein Exoskelett, das Exergames und Assessments zum Üben und Messen der Funktion

der oberen Gliedmassen bietet, werden gegeben. Neue Technologien, insbesondere Robotergeräte in Kombination mit Exergames, können eine grosse Anzahl von Wiederholungen während der gleichen Behandlungssitzung ermöglichen und sind vielversprechend für das motorische Lernen und damit für die Neurorehabilitation. Kapitel 1 beschreibt jedoch auch die Herausforderungen speziell in der pädiatrischen Neurorehabilitation: Es wird langweilig, die gleiche Aufgabe immer und immer wieder zu wiederholen, so dass ein Therapeut zur Verbesserung der Motivation damit beginnen kann, andere Aufgaben in dieselbe Sitzung einzubauen. Hier kommt der kontextuelle Interferenzeffekt ins Spiel. Wenn mehrere motorische Aufgaben während der gleichen Sitzung geübt werden, können sie sich gegenseitig funktionell stören. Wenn die Aufgaben in Übungsblöcken geübt werden (z. B. AAA..., BBB...., CCC....), ist die Interferenz gering, wenn sie in zufälliger Reihenfolge geübt werden (z. B. ABBCABCAC...), ist die Interferenz hoch. Der Effekt der kontextuellen Interferenz und dessen Einfluss auf das motorische Lernen wurde hauptsächlich bei gesunden Erwachsenen untersucht. Hohe kontextuelle Interferenz (zufällige Übungsreihenfolge) führt zu einer besseren motorischen Leistung bei der Retention und dem Transfer im Vergleich zu niedriger kontextueller Interferenz (Übungsblöcke). Während der Acquisition führt das Üben in Blöcken zu einer besseren motorischen Leistung im Vergleich zur zufälligen Übungsreihenfolge. Das Kapitel endet mit der persönlichen Motivation des Autors, diese Dissertation durchzuführen, gefolgt von der Gliederung und Zielsetzung der Arbeit. Das allgemeine Ziel war es, den kontextuellen Interferenzeffekt bei Kindern mit motorischen Beeinträchtigungen aufgrund von Hirnläsionen zu untersuchen.

In **Kapitel 2** wird eine systematische Literaturrecherche und Bewertung der Evidenz über kontextuelle Interferenz bei sich typisch entwickelten Kindern und Kindern mit angeborenen oder erworbenen Hirnläsionen vorgestellt. Die systematische Übersichtsarbeit bietet einen Überblick über 25 Arbeiten, die 27 Experimente zur Bewertung des kontextuellen Interferenzeffekts (Üben in Blöcken versus zufällige Übungsreihenfolge bei sich typisch entwickelnden Kindern) umfassen. Im Allgemeinen war die methodische Qualität der eingeschlossenen Studien niedrig und das Risiko einer Verzerrung hoch. Nur eine Studie untersuchte eine Gruppe von Kindern mit Zerebralparese zusammen mit einer Gruppe von sich typisch entwickelnden Kindern. Die Studie präsentierte die Ergebnisse nur der Gruppe der Kinder mit Zerebralparese nicht. Es wurde eine Best-Evidence-Synthese der Studienergebnisse von sich typisch entwickelnden Kindern durchgeführt. Sie ergab eine begrenzte Evidenz für den Nutzen während der Acquisition vom Üben in Blöcken gegenüber dem Üben in zufälliger Reihenfolge für drei Aufgaben. Analysen der Retention zeigten begrenzte Evidenz für das Üben in zufälliger Reihenfolge gegenüber dem Üben in Blöcken für zwei Aufgaben. Die Analyse der Transferergebnisse ergab für zwei Aufgaben mässig konsistente Evidenz für den Vorteil des Übens in zufälliger Reihenfolge gegenüber dem Üben in Blöcken.

In diesem Kapitel wird der Einfluss von motorischen Aufgaben und -variationen, Erfahrung und Alter der Lernenden sowie der Aufgabenschwierigkeit auf den kontextuellen Interferenzeffekt diskutiert. Ausserdem wird die Frage behandelt, ob die Erkenntnisse über den kontextuellen Interferenzeffekt, die in verwandten Populationen gewonnen wurden (d.h. Erwachsene mit Hirnläsionen, sich typisch entwickelnde Kinder), auf Kinder mit Hirnläsionen übertragen werden können. Zum Beispiel sind die körperlichen Anforderungen im Vergleich zu sich typisch entwickelten Kindern in den anderen beiden Gruppen unterschiedlich. Aufgrund einer Hirnschädigung kann die Fehlererkennung und damit das motorische Lernen beeinträchtigt sein. Sicherlich ist weitere Forschung zu diesem Thema notwendig, um die Auswirkungen von pathologischen Veränderungen auf das motorische Lernen vollständig zu verstehen.

In **Kapitel 3** wird eine Studie über die Reliabilität von und die Erfahrungen mit den Assessments von ChARMin, einem Exoskelett zum Üben und Messen der Funktion der oberen Gliedmassen, vorgestellt. Diese Studie wurde durchgeführt, weil wir planten, ChARMin als Gerät für Kinder mit Hirnläsionen einzusetzen, um verschiedene motorische Aufgaben entweder in Blöcken oder in zufälliger Reihenfolge zu üben.

Stationäre und ambulante Patienten der Kinder Reha Schweiz im Alter von fünf bis 18 Jahren, mit angeborenen oder erworbenen Hirnläsionen, die die Funktion der oberen Gliedmassen beeinträchtigen, nahmen an zwei Terminen teil. Beide Termine beinhalteten die Durchführung derselben Assessments in zufälliger Reihenfolge. Es wurden sieben ChARMin-Assessments durchgeführt. Gemessen wurden der aktive Bewegungsumfang (aROM), der passive Bewegungsumfang (pROM), die Kraft (Strength), der Widerstand gegen passive Bewegungen (RPM), die Qualität der zielgerichteten Bewegungen (QoM), die Tracking-Fähigkeit des Arms (Circle) und der maximal erreichbare Arbeitsbereich (Workspace). Die relative Reliabilität wurde durch die Berechnung von Intraclass-Korrelationskoeffizienten (ICC) ermittelt. Zur Bestimmung der absoluten Reliabilität wurde die Smallest Real Difference als absolutes Verhältnis zum grossen Mittelwert (|SRD%|) berechnet.

Die relative und absolute Reliabilität der Assessments und ihrer einzelnen Parameter variierte stark. Die höchste ICC wurde für den Parameter ‘Maximaler Abstand nach vorne’ des Workspace-Assessments ermittelt (ICC = 0,95, 95% Konfidenzintervall (95%CI) [-0,41, 0,36]). Der niedrigste ICC wurde für den Parameter ‘Widerstand gegen Schulteraussenrotation’ des RPM-Assessments ermittelt (ICC =- 0,03, 95%CI [-0,41, 0,36]). Der niedrigste |SRD%| wurde für den pROM-Parameter ‘Schulterextension’ ermittelt (5,9%), während der höchste Wert für den RPM-Parameter ‘Widerstand gegen die Unterarmpronation’ registriert wurde (41’810,1%).

Innerhalb dieses Kapitels wurden die Ergebnisse diskutiert und mit der vorhandenen Literatur verglichen. Zusätzlich zu den Reliabilitätsergebnissen wurden bei der

Anwendung dieser Assessments einige wichtige Punkte aufgedeckt, die in unseren Empfehlungen für den klinischen Einsatz der Assessments berücksichtigt wurden. Es wurde festgestellt, dass die Anwendung von aROM und pROM nicht empfohlen wird, da der Bewegungsumfang durch das Gerät aufgrund von Sicherheitsvorschriften mechanisch begrenzt ist. Auch RPM wird aufgrund der durchweg schlechten Reliabilitätsresultaten und der Unsicherheiten darüber, welches Konstrukt RPM tatsächlich misst, nicht empfohlen. Strength zeigt Potenzial für den klinischen Einsatz, jedoch wird empfohlen, das Messverfahren und die Datenauswertung anzupassen. Wir empfehlen die Verwendung der Assessments QoM, Circle und Workspace in der klinischen Praxis. Die meisten ihrer Parameter zeigten eine akzeptable Reliabilität und es gibt keine alternativen konventionellen Assessments, die ähnliche Parameter bewerten. Dennoch ist Vorsicht geboten, da schwerer beeinträchtigte Kinder Schwierigkeiten bei der Durchführung der QoM- und Circle-Assessments hatten.

In **Kapitel 4** werden die ersten Durchführbarkeitsexperimente mit ChARMin vorgestellt, die das Üben von Exergames in Blöcken und in zufälliger Reihenfolge vergleichen. Es handelte sich um einen begrenzten und überschaubaren Ansatz mit dem Hauptziel, einige Informationen über die Akzeptanz der Teilnehmer, die Implementierung der Intervention und die Praktikabilität eines zukünftigen Studienablaufs zu gewinnen. Kinder mit angeborenen oder erworbenen Hirnläsionen, die sich in der Neurorehabilitation befanden, wurden rekrutiert, um an zwei Terminen zwei verschiedene Exergames entweder in Blöcken oder in zufälliger Reihenfolge zu üben. Der erste Termin war während der vorletzten Therapiesitzung des Tages geplant, der zweite Termin während der ersten Therapiesitzung am nächsten Morgen. Vor und nach beiden Sitzungen und eine Stunde nach der ersten Sitzung wurden Assessment-Blöcke (Box and Blocks Test und die beiden ChARMin Assessments QoM und Circle) durchgeführt. Wir kategorisierten die Notizen, die während des Studienablaufs gemacht wurden, in ‘Akzeptanz der Teilnehmer’, ‘Umsetzung der Intervention im klinischen Kontext’ und ‘Praktikabilität des Studienablaufs’. Die Hauptprobleme des Verfahrens waren die technischen Probleme, die wir mit dem ChARMin-Gerät erfuhren (z.B. Fehler, die zu einer Unterbrechung der Anwendung des Gerätes führten). Anzahl und Schwere dieser Probleme schienen eng mit der Compliance der Teilnehmer zusammenzuhängen. Diese Ergebnisse und weitere Aspekte der Durchführung der Studie und ihrer einzelnen Komponenten werden diskutiert. Am Ende des Kapitels kamen wir zu dem Schluss, dass umfangreiche Anpassungen erforderlich sind, um eine wissenschaftlich fundiertere Studie, die auf dieser Studie aufbaut, fortzuführen. Daher haben wir zunächst gemeinsam mit dem Ingenieur von ChARMin Anpassungen und Fehlerbehebungen diskutiert. Basierend auf den Ergebnissen dieses Experiments entwickelten wir auch ein detailliertes Studienprotokoll für eine Pilotstudie über den kontextuellen Interferenz-Effekt bei Kindern mit angeborenen oder erworbenen Hirnläsionen, die robotische Exergames entweder in Blöcken oder in zufälliger Reihenfolge üben.

In **Kapitel 5** wird das oben erwähnte detaillierte Studienprotokoll einer Pilotstudie beschrieben, die die das Üben in Blöcken versus Üben in zufälliger Reihenfolge eines Robotik-Exergames für die oberen Gliedmassen untersucht. Dieses Kapitel weist darauf hin, dass das Wissen über den kontextuellen Interferenzeffekt bei Kindern mit angeborenen und erworbenen Hirnläsionen von Interesse ist, um die Ergebnisse der Neurorehabilitation bei diesen Kindern potenziell zu verbessern. In diesem Kapitel stellen wir das Protokoll einer Pilotstudie vor, indem wir zehn Fragen zur Machbarkeit evaluieren:

1. Sind die Einschlusskriterien spezifisch genug, um eine Stichprobe von Teilnehmern zu erhalten, die für diese Pilotstudie geeignet ist (z.B. sind sie in der Lage, das Exergame zu spielen)?
2. Ist die Rekrutierungsrate praktikabel?
3. Ist der Ablauf der Planung machbar?
4. Ist das Randomisierungsverfahren praktikabel?
5. Ist das gesamte Verfahren für die Teilnehmer machbar?
6. Ist die Handhabung des Roboters praktikabel?
7. Ist der Umgang mit der grossen Datenmenge machbar?
8. Sind die Outcomes Messverfahren in diesem Setting responsiv und sensitiv genug?
9. Gibt es einen Einfluss von parallel zur Studie stattfindenden Therapien innerhalb des Reha-Settings auf die Studienergebnisse?
10. Ist die Durchführung der Hauptstudie im Hinblick auf die benötigte Stichprobengrösse, die aus den Daten für den primären motorischen Lernerfolg berechnet wurde, machbar?

Wir wählten ein Exergame, das mit Bewegungen der oberen Gliedmassen in der frontalen oder horizontalen Ebene gespielt werden kann. Der/die Spieler/in muss Greifbewegungen in verschiedene Richtungen ausführen, um acht Ziele pro Versuch zu erreichen, die radial um einen Mittelpunkt erscheinen. Durch das Üben dieser Variationen (d.h. das Bewegen in verschiedenen Ebenen) des Exergames sollte ein unterschiedliches Mass an kontextueller Interferenz induziert werden (d.h. niedriges Mass an kontextueller Interferenz durch Üben in Blöcken, hohes Mass an kontextueller Interferenz durch zufällige Übungsreihenfolge). Es wurde ein dreiwöchiger Studienablauf geplant: In Woche 1 (Kontrollwoche) waren zwei Assessmentblöcke geplant. Woche 2 (Übungswoche) bestand aus drei Übungssitzungen, in denen die Teilnehmer jeweils 30 Versuche mit jeder Version des Exergames durchführen sollten. Unmittelbar vor der ersten Übungseinheit, unmittelbar nach, eine Stunde nach,

einen Tag nach und eine Woche nach der letzten Übungseinheit sollten Assessments durchgeführt werden. Folglich diente Woche 3 als Follow-up-Woche.

Kapitel 5 schliesst mit einer Diskussion über den weiteren Kontext, in den die Pilotstudie gestellt werden kann: Sie ist ein erster Schritt, um mehr Informationen über den kontextuellen Interferenzeffekt bei Kindern mit angeborenen oder erworbenen Hirnläsionen zu erhalten.

Kapitel 6 beinhaltet die Ergebnisse der Pilotstudie, deren Protokoll in Kapitel 5 beschrieben ist, vorgestellt. Wir rekrutierten von den stationären Patienten/innen der Kinder Reha Schweiz Teilnehmer/innen während eines Jahres. Die Reihenfolge der zehn Fragen zur Machbarkeit wurde entsprechend der Wichtigkeit der Fragen angepasst (die erste Frage ist die wichtigste). Die Analyse der Daten zu den zehn Fragen führte zu folgenden Ergebnissen:

1. Die Ein- und Ausschlusskriterien waren nicht spezifisch genug, da zwei Teilnehmer/innen körperlich nicht in der Lage waren, die Exergames zu spielen.
2. Rekrutierungsrate: Wir informierten 18 in Frage kommende Kinder über die Studie, 12 von ihnen stimmten einer Teilnahme zu und wurden eingeschlossen. Es gab aus verschiedenen Gründen mehrere Drop-outs in verschiedenen Phasen der Studie. Am Ende standen vier vollständige Datensätze und zwei unvollständige Datensätze (fehlende Follow-up Messung und ausgeschlossene Daten von Woche 1) zur Verfügung. Diese Rate ist zu niedrig und nicht praktikabel.
3. Eine Berechnung des Stichprobenumfangs auf Basis der erhobenen Daten dieser Studie ergab, dass eine zukünftige Hauptstudie 15'355 Proband/innen erfordern würde, und daher nicht durchführbar wäre.
4. Für 50% der Teilnehmer war das Verfahren nicht durchführbar, da sie nicht in der Lage waren, 30 Versuche jedes Exergames während jeder Sitzung durchzuführen.
5. Die Planung der Studientermine war schwierig, vor Allem weil sie zwischen den regulären Therapieterminen der einzelnen Kinder geplant werden mussten. Kurzfristig geplante obligatorische Reha-Termine oder Krankheit eines Teilnehmers waren zusätzliche Hindernisse.
6. Das Randomisierungsverfahren funktionierte gut, die Randomisierung nach dem Minimierungsprogramm konnte problemlos angewendet werden.
7. Die Handhabung des Roboters erwies sich im Vergleich zu den ersten Machbarkeitsexperimenten (Kapitel 4) als besser, dennoch gab es immer noch technische Probleme, die die Daten wahrscheinlich beeinflussten.
8. Die Handhabung der Datenmenge war machbar, jedoch würde eine automatisierte Methode zur Analyse der ChARMin-Daten das Verfahren erleichtern.

9. Die Outcome Messverfahren waren nur schwach responsiv, dennoch müssen diese Ergebnisse mit Vorsicht behandelt werden, da sie auf der Grundlage einer begrenzten Datenmenge berechnet wurden.
10. Parallel stattfindende Therapien schienen die Ergebnisse nicht zu beeinflussen. Doch auch hier ist die zugrunde liegende Datenmenge begrenzt.

Kapitel 6 diskutiert die Ergebnisse der Machbarkeitsstudie und vergleicht sie mit der vorhandenen Literatur. Zusammenfassend lässt sich sagen, dass eine Hauptstudie mit dem Vorgehen dieser Pilotstudie nicht durchführbar ist. Vielleicht wäre ein alternatives Studiendesign sinnvoller, um Informationen über den kontextuellen Interferenzeffekt bei Kindern mit Hirnläsionen zu sammeln.

In **Kapitel 7**, der allgemeinen Diskussion dieser Dissertation, werden die Hauptergebnisse und die kritische Diskussion spezifischer Aspekte der in dieser Dissertation enthaltenen Studien erörtert. Das Hauptaugenmerk liegt dabei auf dem Prozess der Machbarkeits- und Pilotstudien. Je nach Literatur werden die Bezeichnungen “Machbarkeitsstudie” und “Pilotstudie” austauschbar verwendet oder manchmal mit unterschiedlichen Bedeutungen versehen: Im Vergleich zu einer Machbarkeitsstudie hat eine Pilotstudie ein strengeres Protokoll und eine Absicht für zukünftige Arbeiten, nämlich eine Hauptstudie.

Es wurde auch vorgeschlagen, die Fragen “Kann es funktionieren?”, “Funktioniert es?”, “Wird es funktionieren?” zu verwenden, um einen Machbarkeits-/ Pilotstudienprozess zu planen und zu präsentieren. Die in Kapitel 4 beschriebene Vorläuferstudie (“Kann es funktionieren?”) und die in den Kapiteln 5 und 6 beschriebene Pilotstudie (“Funktioniert es?”) werden in den Rahmen dieser Fragen eingefügt. Kapitel 7 schliesst mit einem Abschnitt über Implikationen für zukünftige Forschung ab. Basierend auf den Erfahrungen, die wir in diesem Projekt gemacht haben, ist es fraglich, ob eine randomisierte kontrollierte Wirksamkeitsstudie tatsächlich geeignet ist, den kontextuellen Interferenz-Effekt in der pädiatrischen Neurorehabilitation zu untersuchen. Diese Frage wird diskutiert und es werden Vorschläge für alternative Studiendesigns gemacht, welche die Individualität der Teilnehmer mehr berücksichtigen (z.B. Studien im Single-Subject-Design, Studien im Mixed-Methods-Design). Zusätzlich wird betont, dass die wichtigen Stakeholder (nämlich Vertreter der Population von Kindern in der Neurorehabilitation, deren Eltern, Geschwister, Freunde und Betreuer) mehr einbezogen werden sollten, um die Relevanz der Forschung zu erhöhen.

IMPACT

INTRODUCTION

The Cambridge Dictionary defines “impact” as “a powerful effect that something, especially something new, has on a situation or person” [1]. Concerning this doctoral thesis, “something new” can be seen as the results we obtained with the research that was performed. The nature and strength of the effect our research, in which situations the effect comes into play and which persons are involved, is discussed in the follwoing paragraphs.

THE RELEVANCE OF THIS RESEARCH

According to Article 32 of the Swiss Federal Health Insurance Act, benefits covered by compulsory health insurance need to be economical, effective and expedient [2]. To ensure optimal treatment of patients, in this case the treatment of the young persons with neurological disorders, evaluating the effectiveness of therapies is needed. Hence, investigating the potential benefit due to contextual interference on motor learning is meaningful. Our systematic review brought forth, that there is a knowledge gap when it comes to the effect of contextual interference in children with brain lesions.

New technologies can seem promising in terms of effectiveness and are increasingly applied in the rehabilitation setting. Yet, it is still not completely understood how these technologies should be implemented best as an integral part of rehabilitation therapies. Implementing new technologies in research is of interest and relevance too as this provides knowledge in this rather young field. We used the ChARMin robot for our projects and thus aimed to gain more knowledge about the use of robotic devices in research. When applying any assessment, be it a conventional assessment or provided by new technologies, it is important to know the clinimetric properties (i.e. reliability, validity, and responsiveness) for the particular population within the context of the clinical application (i.e. for diagnostics, evaluation or prediction). Clinimetric properties are quality characteristics of an assessment. Having knowledge about the clinimetric properties of an assessment is significant as it puts the effort and benefit ratio into perspective for the user.

We assessed the reliability of assessments measuring upper limb functions provided by ChARMin. Reliability is an important clinimetric ability referring to the repeatability and consistency of the results recorded with the assessment. A necessary and important step within the research process. However, just like all clinimetric properties, reliability also depends on the population as well as the context in which the assessments are applied. Therefore, we might have provided important information that can be applied to our setting, with the children undergoing

neurorehabilitation, yet, this information is not transferable to other populations or settings. Hence, the relevance of our reliability results is limited.

As economical aspects are not only an issue in the clinical area but also in research, a careful consideration of the limited resources is required. Research is mainly financed by third party money and justification for spending these funds is required. Hence, it is of relevance to assess whether conducting an effectiveness study requiring a lot of resources is actually feasible.

WHO BENEFITS FROM THIS RESEARCH?

An important aspect of the impact of research, are the persons or groups of persons deriving advantage or benefit from this research. The relevance that has been evaluated in the previous paragraph and the benefit are dependent on each other as the relevance is determined by these very same people. So who benefits from this research?

The objectives of this doctoral thesis that were defined in the beginning, focused on the improvement of therapy schedules. We expected to receive and transfer the knowledge about a higher benefit of one practice order over the other one to clinical practice. Yet, what we gained is knowledge about processes needed before even starting with a large effectiveness study on that goal. Initially, we expected that we could present an optimised therapy schedule based on findings about whether blocked or random practice would be beneficiary for the neuropaediatric population. Hence, children with brain lesions undergoing neurorehabilitation would be beneficiaries in terms of them profiting from this improved schedule with better motor learning outcomes. Additionally, their therapists would benefit in terms of gained knowledge about how to organise a practice session to attain an optimised learning situation. Yet, as it turned out, the benefits of this research were distributed differently, this is discussed in the following paragraphs.

The research community

One group of persons potentially deriving advantage from research is the research community. Reading studies published by peers leads to an increase in one's own knowledge, be it methodological or content-related. Promoting pilot studies prevents researchers and funders from spending resources to studies doomed to fail: "[…] it is worth making clear that a pilot study that shows the main study is not likely to be feasible is not a failed (pilot) study. In fact, it is a success- because you avoided wasting scarce resources on a study destined for failure!" p. 6 [3] [2]. Hence, findings of our pilot study prevented our research group from spending third party money on a project that turned out not feasible, and also protected a vulnerable population of children with brain lesions from attending a study without prospective gains. Yet, our findings contributed to the body of knowledge in the field of methodology of pilot

studies. This could promote the conduct and publication of pilot studies in our field. Our systematic review showed that studies assessing the contextual interference effect in children were of very low methodological quality. Publishing study protocols in peer reviewed journals exposes elaborated research ideas to an audience of experts, before the actual study is conducted. This is an important step to make full research processes more transparent and therewith increase the quality of research. The conclusion of this dissertation is that it is valuable to look outside the box of conventional research approaches in paediatric neurorehabilitation. The rather small field, the heterogeneous population, the individual presentation of disease patterns or levels of severity and disability makes it difficult to design and follow the strict rules of a randomised controlled effectiveness trial needed to attain a sound methodology. More preparation including a feasibility randomised controlled trial as part of the whole trajectory towards an effectiveness randomised controlled trial is urgently needed. This has implications for future research, namely, increased applications of more and also different research approaches in paediatric neurorehabilitation. For example, more individual approaches could probably contribute to make the research more meaningful for our young participants and to be able to draw valid conclusions. First, the awareness and then the promotion of more high quality scientific approaches will be of great importance in our field.

The developers of robotic devices

Robotic therapies and assessments became popular in the last decades. They provide certain advantages compared to conventional therapies (e.g. number of repetitions, intensity, highly standardised measurements). Yet, despite these advantages and the appealing futuristic appearance of new technologies, it is important to look at what is behind. Clinical and research experience with ChARMin is not yet extensive. With the results obtained with the reliability investigation of the ChARMin assessments, we added information that will be used to improve and further develop the measurement of upper limb functions with ChARMin. As an example, we recommended to adapt the strength assessment in such a way that the instructions for the participants would be easier to understand and follow. This would improve reliability of the strength measurements. This recommendation has been well received and partly already been implemented for a future update of the ChARMin software. However, as ChARMin is the only device of its kind, these updates will only affect our centre and no immediate generalisation can be made.

Persons involved in clinical practice: children and therapists

Children with affected upper limb function undergoing neurorehabilitation in our centre still profit to some extent from the results obtained from our reliability study. Assessments were and will be adapted according to our experience, most likely leading to more reliable outcomes; others were not applied in studies as they showed low reliability and adaptations were not possible. In the clinical setting, therapists

can for example evaluate whether changes measured with ChARMin assessments actually reflect a real change and are not only due to a measurement error when comparing them with the results of the absolute reliability measures (i.e. smallest real difference). Knowing about a true change ensures children that the training is beneficial for them and potentially increases their motivation.

In the course of the pilot study, new exergames have been programmed, increasing the range of robot exergames for practicing the function of the upper limbs in our centre. These innovations could provide for more motivation, variety and diversity in training. Again, for now, this benefit only applies to our centre.

KNOWLEDGE TRANSFER

Spreading knowledge and making it accessible and usable is a fundamental part of the management of knowledge [4] [3]. Knowledge without transfer is an artefact. Knowledge can be transferred to different kinds of receivers with various media. The scientific findings from this thesis have partly been published in scientific journals and will be published in the future. Knowledge addressing the interested scientific community thus will be transmitted via relevant journals. Yet, optimally, transfer would also include the use or implementation of the according knowledge. This again requires the receiver to absorb the provided information and apply it in his work field.

Knowledge addressing the developer of the robotic device has already been transferred. Direct contact with the developer about the issues of the robotic device and its software led to several adaptations already during this doctoral thesis. Improvements were implemented due to the findings of not only the reliability study but also due to the experiences we made during the feasibility and pilot study process. Further adaptations are planned and will be implemented shortly. Hence, within our centre, this part of knowledge has not only been transmitted but also been transferred.

Knowledge about the contextual interference effect as well as about difficulties of conducting an effectiveness study obtained in the course of this thesis has also been spread within our centre, at national and international congresses. Findings were presented on posters and with oral presentations. Hence, direct contact and discussions with therapists, doctors and interested researchers have taken place. Yet, also here, with this approach, there is limited influence on the knowledge implementation. The most impact could take place within the research community from both, inside and outside our centre, by promoting, that more individual and alternative research approaches could be considered in the very heterogeneous field of paediatric neurorehabilitation.

The big challenge of knowledge transfer and implementation in this case is, that

the results we obtained are mostly limited to our centre. The implementation of the knowledge we gained about the content as well as the methodology will need a long and complex way to go to be implemented in a broader context. Intensive implementation strategies are needed to reach that goal.

CONCLUSION

The research that has been conducted in the context of this doctoral thesis, is relevant for both, research and clinical practice, but mainly in our centre for the moment. Our research group, the research community, developers of robotic devices as well as therapists and children undergoing neurorehabilitation, benefit to a greater or lesser extent. In this thesis we pointed out directions for future research, and that more individual approaches (e.g. single subject design studies, mixed-methods design studies) could be indicated in this heterogeneous population. Reducing the burden for the participants and their families in general as only a fractional part of participants is needed for a study, would be one advantage. Most importantly, the inclusion of children, their parents and other important stakeholders in the research process, would most certainly increase the relevance of the research and therewith its impact.

REFERENCES

1. Cambridge dictionary. Impact. 2021. Dictionary Merriam Webster. Congenital. 2020. https://dictionary.cambridge.org/dictionary/english/impact. Accessed 21 Jan 2021.
2. Die Bundesversammlung der Schweizerischen Eidgenossenschaft. Bundesgesetz über die Krankenversicherung. 2. Abschnitt, Artikel 32. Switzerland; 1994.
3. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: The what, why and how. BMC Med Res Methodol. 2010;10.
4. Paulin D, Suneson K. Knowledge transfer, knowledge sharing and knowledge barriers-three blurry terms in KM. Electron J Knowl Manag. 2012;10:81–91.

ACKNOWLEDGEMENT

During the time I spent doing my PhD, I learnt a lot: From research methodology to neuroscience and how to put little plastic kings into all kinds of food. I drank huge amounts of coffee, a little bit less of wine (the number of glasses highly correlates with the time I spent on my PhD, yet, causality is unknown). Many people accompanied me on my way, inspired and supported me.

Rob, I met you during my Master's and already then I was deeply impressed by your calmness and clear view on all kinds of tricky research issues. Your way of seeing the centre point of a problem helped me more than once to get back on track. I am truly grateful that I had the opportunity to benefit from it and your expertise also during this PhD. I will always remember you saying: "sometimes you just need to start" when we discussed my extensive pre-study planning of the pilot study. Thank you for this push and many other things.

Carolien, your inputs, remarks and questions were a true enhancement of my work during this PhD. I appreciate your kind and empathetic way with which you came across me from the beginning. It was a great and wonderful opportunity to work with you and to benefit from your rich experience in research methodology. During our meetings you often put things into perspective, reinforced my ideas or made quite clear that I should change my way when I was in danger of losing myself in the confusion of science (which was quite often the case). Thank you very much.

Judith, my namesake, thank you for being there during the PhD meetings. Your inputs influenced my work and gave me clues to issues I thought were difficult to resolve. Thank you very much for being a part of the Maastricht-Team.

Huub, I am truly lucky to have you as my supervisor and my boss. I very much appreciate your problem-solving strategies. You always found a solution for all the different kinds of issues, be it employment related or answering a tricky question from a reviewer. You always stayed calm when I came rushing into your office with the typical symptoms of pre-submission panic (which are gasping, ruffled hair, jerky movements, speaking way too fast, panic facial expression, and bringing along a list with many questions, highlighted with all the colours of the rainbow (50% of the questions have been discussed earlier, but just in case, you know). Thank you very much for supporting me through this PhD.

Peter, you were a member of the committee needed for me to attend the PhD programme at the Centre for the Neurosciences at Zurich University and Federal Institute of Technology, Zurich, Switzerland. We did not meet often, but you provided helpful feedback to my yearly progress reports and always in due time. I appreciated this very much. Thank you!

Andreas, thank you very much for leading the Swiss Children`s Rehab, the way you do. It feels like family and to a large deal this is your merit. I had the great opportunity and pleasure to accompany you to two congresses, one to Canada and one to England. Two events, I would consider important on my journey as it enabled me to talk in front of experts about my experiences as a physiotherapist in robotics. Additionally, I will never forget you swinging the napkin.

PRRG, PRRD or however I should call you, I enjoy working with the best team! I never experienced anything alike in the past. Thank you for interesting and fruitful discussions, questions, jokes and empathy in so many situations. Thank each and every one of you for being you and letting me be me.

I am thankful for the financial support by the **Mäxi Stiftung** and the **Clinical Research Priority Program Neurorehabilitation** of the University of Zurich, making my PhD thesis possible.

To all the members of the reading committee and the corona: **Prof. dr. F. J. M. Feron, Prof. dr R. J. Vermeulen, Prof. dr. A. Buizer, Dr. E. A. A. Rameckers, Dr. B. Tobler-Ammann, Prof. dr. K. Klingels**, thank you all for your time you spent reading and evaluating my thesis, your interest in my work and your questions.

Corinne, I first met you at my former work place, where you were recruiting participants for your Lokomat trial together with Andreas. I was so impressed, meeting a physiotherapist working in research, what glory! Indeed, I am still and even more impressed. Your calm and still enthusiastic way to discuss things that are important to you and your sophisticated point of view on topics of daily life, are only two things of many, I could mention here. Thank you, for your multifaceted support, for being one of my paranymphs, and your friendship.

Petra, Peti, Marpe, Frau Maricek or Marisco – I guess you won the title of the person with the most nicknames in the team. We met during the Master's and started off with a huge success in the first paediatric physiotherapy teamwork. You were one of the reasons I started working at the Swiss Children's Rehab. Thank you for your down-to-earth nature. I appreciate your advice both in a professional and personal sense, depending on my needs they make me dream or bring me back to reality. I am looking forward to further joint projects. Reach for the stars!

Team ChARMin: **Anja, Jan** and **Urs**, you are the best sub-team I could have dreamt of. You supported me with your technical and clinical knowledge, funny and challenging picture puzzles, good talks – professional and private, and chocolate. Our joint-work was and is incomparable, thank you from the bottom of my heart. Team ChARMin forever!

Laura, thank you very much for working with me! Looking something up in the notes of your measurements made me smile more than once when noticing your charming Ticino accent even in the written words. It always felt like a sun coming up.

Monica, you were my first Master student I was supervising. We were a wonderful team. Your thorough and well-thought way to tackle your project was inspiring. I appreciate your support very much!

To **all the children** participating in our research projects and **their parents**: thank you very much for your time, your efforts and interest. We appreciate your contribution very much!

Simone, my simultaneous Maastricht-PhD-colleague, we wrote so many e-mails hence and forth, most of them about " I Weiss gar ned wieni wiiter chome, wie gohts der?". They were always a great relieve because I could see that it seemed to be quite normal to struggle sometimes. I was always looking forward to seeing your friendly face at the PhD-meetings in Winterthur and Sierre and discussing issues (of any kind) with you. Thank you for being my paranymph at the defense.

Verena and Felix, my wonderful parents, thank you very much for supporting me all this time. You accompanied me through so many steps in my life, some of them were not easy for neither of us. You always believe(d) in me and my ability to decide myself what is best for me. This is indeed not self-evident. I Am truly lucky to have the two of you as my parents. Especially, I want to thank you for supporting our family by looking after Ina each and every week. She enjoys the time she spends with you very much.

Ingrid and Adri, thank you both for looking after Ina whenever I needed additional time for my work. She loves to be with you because she is allowed to do things she cannot do at home. I appreciate your support immensely!

Mirjam, Jonas and Tobias, best siblings ever! The openness and directness inherent in our being together are incomparable and helped me through some very tough times. I thank you for this with all my heart! You all are a big and immensely important part of my journey. But no, and again, I will not be allowed to issue prescriptions, DIFFERENT KIND OF DOCTOR! A very special thank goes to Mirjam, supporting me with her talent and illustrative view by giving me advice for the layout, the cover and the presentation of this thesis.

Martina, you are not a part of the research team, not even a physiotherapist. We met at the Kanti, in 1996, when we both were young, foolish and crispy. No, we may not be young anymore, we still are kind of foolish though, and most certainly, we still are crispy (I get audio evidence every morning, when trying to get out of bed). You were my first friend with a PhD, I admire(d) you, and I referred to you as my "PhD-friend". Yet, you are so much more, there are so many other even more inspiring facts about you, my Fründin. Thank you for being you and a part of my life.

Ina, you remind me of the importance of childlike curiosity every day. You surprise me. You leave me in utter astonishment. You make my life so colourful (the whole spectrum). And probably most important of all, you challenge me. Being with you

sets so many things in another perspective. Thank you for being so funny, loving and stubborn, I learn a lot from you.

Reto, you asked me once: “How is it possible to be bright and so stupid at the same time?” The most wonderful (the “bright”-part) and true (the “stupid”-part) sentence anyone ever said to me. I was going through a rather destructive phase of self-doubt at the time and this question brought me back to reality. I appreciate your last-minute support with the layout of posters and especially this dissertation so very much, your mental last-minute support is priceless. I love the life I am sharing with you, the wonderful days with the “Büssli”, your support, your critical questions, the discussions we usually have in the evenings in the kitchen, your understanding, and you. Thank you from the bottom of my heart for all of this and so much more.

ABOUT THE AUTHOR

Judith Verena Graser was born on March 24, 1980 in Aarau, Switzerland. She earned her diploma as a physiotherapist in 2000 at the School of Physiotherapy, Aargau, Schinznach, Switzerland.

Her professional experience began at the Vivendra Foundation in Dielsdorf, Switzerland, where Judith worked with children, adolescents and adults with cerebral palsy. Her main professional interest has always been in the neuropaediatric field and after completing the Bobath introductory course for adult patients in Zurzach, Switzerland in 2006, she attended the paediatric Bobath training at the Bobath Centre in London, UK in 2008.

Judith's interest in evidence-based therapy and desire for a career change led to the decision to pursue a master's degree in physiotherapy at the Zurich University of Applied Sciences, Winterthur, Switzerland. In addition to studying part-time, Judith began working as a physiotherapist on the robotics team and as a research assistant at Swiss Children's Rehab, University Children's Hospital, Zurich, Switzerland, in 2011. Judith completed her master's degree in 2013 with a thesis on the reliability of measuring postural control with inertia measurement units in healthy adults and patients with chronic low back pain.

While the topic of the thesis was chosen according to the situation and opportunities at the time, Judith's main research interest remained in the field of neuropaediatrics. In 2015, she started her PhD as an external doctoral candidate at the Faculty of Health, Medicine and Life Sciences at Maastricht University, Maastricht, The Netherlands. After the birth of her daughter in 2017, she resigned from her position in the robotics team to focus on her PhD research. She conducted the research described in this thesis under the supervision of Prof. Dr. Rob de Bie, Dr. Caroline Bastiaenen, and Prof. Dr. Hubertus van Hedel.

Judith is a member of the international editorial board of the journal Pediatric Physical Therapy and a board member of the Swiss Association of Pediatric Physical Therapists.

LIST OF PUBLICATIONS

Aurich (-Schuler) T, Warken B, **Graser JV**, Ulrich T, Borggraefe I, Heinen F, et al. Practical recommendations for robot-assisted treadmill therapy (Lokomat) in children with cerebral palsy: indications, goal setting, and clinical implementation within the WHO-ICF framework. Neuropediatrics. 2015;46:248–60.

Graser JV, Letsch C, van Hedel HJA. Reliability of timed walking tests and temporo-spatial gait parameters in youths with neurological gait disorders. BMC Neurol. 2016;16:1–12.

Rast FM, **Graser JV**, Meichtry A, Ernst MJ, Bauer CM. Reproducibility of a new signal processing technique to assess joint sway during standing. J Biomech. 2017;51:133–6.

Graser JV. Spezifische Roboteransätze für Kinder. neuroreha. 2017;09:172–8.

Van Hedel HJA, Severini G, Scarton A, O'Brien A, Reed T, Gaebler-Spira D, […], **Graser JV**, et al. Advanced Robotic Therapy Integrated Centers (ARTIC): An international collaboration facilitating the application of rehabilitation technologies. J Neuroeng Rehabil. 2018;15.

Graser JV, Bastiaenen CH, van Hedel HJA. The role of the practice order: a systematic review about contextual interference in children. PLOS One. 2019;14.

Graser JV, Bastiaenen CHG, Keller U, van Hedel HJA. Contextual interference in children with brain lesions: protocol of a pilot study investigating blocked vs. random practice order of an upper limb robotic exergame. Pilot Feasibility Stud. 2020;6.

Graser JV, Bastiaenen CHG, Gut A, Keller U, van Hedel HJA. Contextual interference in children with brain lesions: a pilot study investigating blocked vs. random practice order of an upper limb robotic exergame. Under review.

Graser JV, Prospero L, Liesch M, Keller U, van Hedel HJ. Reliability of and practical lessons learned from robotic upper limb assessments in children undergoing neurorehabilitation. Submitted.