Does the use of home-based assistive rehabilitation technology enhance the functional benefits of botulinum toxin in children with cerebral palsy who have upper limb movement difficulties: a single-blind randomised controlled trial.

Nicholas John Preston

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Academic Department of Rehabilitation Medicine

October 2014

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Further details of the jointly-authored publications and the contributions of the candidate and the other authors to the work should be included below this statement.

Acknowledgements

Contributions to the research described in the thesis

This research described in this thesis has been carried out by a team which has included Professor Bipin Bhakta, Professor Martin Levesley, Professor Mark Mon-Williams, Dr Rory O'Connor, Dr Andrew Weightman, Dr Mike Clarke, Dr Ray Holt, Dr Matthew Allsop, Mr Justin Gallagher and Mr M.C. Horton. My own contributions, fully and explicitly indicated in the thesis, have been the design, management and conduct, data collection and data analysis of the study to develop the Children's Arm Rehabilitation Measure and the randomised control trial with supervision by Professor Bhakta, Professor Martin Levesley, Professor Mark Mon-Williams and Dr Rory O'Connor. Other research described in the thesis includes my own contributions and that of other members of the group:

Chapter 2, sub-paragraph 2.2: Design, construction and feasibility trial of prototype home-based assistive joystick and computer games (NIHR-funded study, grant ID G006). WEIGHTMAN, A. P. H., PRESTON, N., HOLT, R., ALLSOP, M., LEVESLEY, M. & BHAKTA, B. 2009. Engaging children in healthcare technology design: developing rehabilitation technology for children with cerebral palsy. *Journal of Engineering Design,* 21**,** 579-600.

I was responsible for the identification of schools, recruitment of children, taking of informed consent and was heavily involved in the user-centred design procedures. I was responsible for the day-to-day conduct of the feasibility trial, delivering the computer games to children's homes and carrying out baseline and follow up functional assessments. I supported the researchers who captured the kinematics of the children's upper limb. I carried out the data analysis of functional measures and prepared the data for publication. Dr Weightman constructed the joysticks and wrote all software for games and for the assistive guidance that powered the joysticks. Dr Weightman was involved in the design procedures, supported the day-to-day running of the trial and carried out the kinematic assessments. Dr Weightman wrote the publication, which was reviewed and amended as necessary by all other authors. Dr Allsop and Dr Holt were primarily responsible for the user-centred design procedures. Professor Levesley and Professor Bhakta and had overall responsibility for the study, and were members of the steering committee, supervising all aspects of the study.

Chapter 2, sub-paragraph 2.3: Design, construction and feasibility study of school-based computer-assisted arm rehabilitation games system for children with cerebral palsy (NIHR-funded study, grant ID K005). PRESTON, N., WEIGHTMAN, A., GALLAGHER, J., HOLT, R., CLARKE, M., MON-WILLIAMS, M., LEVESLEY, M. & BHAKTA, B. 2014. Feasibility of school-based computer-assisted robotic gaming technology for upper limb rehabilitation of children with cerebral palsy. *Disability and Rehabilitation: Assistive Technology,* 0**,** 1-8.

I was responsible for the approach to participating schools and centres, recruitment of children, taking of informed consent and was involved in the user-centred design procedures. I was responsible for the day-to-day conduct of the trial, delivering the computer games to children's schools and carrying out baseline and follow up functional assessments. I captured the kinematics of the children's upper limb using a novel, experimental device. I carried out the data analysis of functional measures, of the kinematics and prepared the data for publication. Dr Weightman constructed the assistive gaming devices and wrote all software for games. Mr Gallagher was responsible for the assistive software that powered the robotic arm. Dr Weightman was involved in the design procedures and responsible for the overall conduct of the study. I wrote the publication, which was reviewed and amended as necessary by all other authors. Dr Holt was primarily responsible for the user-centred design procedures. Professor Levesley, Professor Mon-Williams and Professor Bhakta had overall responsibility for the study, and were members of the steering committee, supervising all aspects of the study. Professor Levesley supported delivery of games to schools. Dr Clarke was the clinical supervisor and part of the steering committee.

Chapter 3, sub-paragraph 3.6 Assessment of upper limb kinematics of children with cerebral palsy: CPKAT.

PRESTON, N., WEIGHTMAN, A., CULMER, P., LEVESLEY, M., BHAKTA, B. & MON-WILLIAMS, M. 2014. The Cerebral Palsy Kinematic Assessment Tool (CPKAT): feasibility testing of a new portable tool for the objective evaluation of upper limb kinematics in children with cerebral palsy in the non-laboratory setting. *Disabil Rehabil Assist Technol***,** 1-6.

This was a sub-study within the study described in the chapter and publication above (Chapter 2, sub-paragraph 2.3: Design, construction and feasibility study of school-based computer-assisted arm rehabilitation games system for children with cerebral palsy (NIHR-funded study, grant ID K005)). I was responsible for all aspects of the conduct of this sub-study, which involved kinematic testing on children recruited to the main study. I collected, organised and performed statistical analysis on the data, and I wrote the publication,

which was reviewed and amended as necessary by all other authors. Professor Levesley and Professor Bhakta and had overall responsibility for the study, and were members of the steering committee, supervising all aspects of the study. Professor Mon-Williams is the joint patent holder who conceptualised CPKAT, and the primary supervisor for the production of the publication. Dr Culmer is the engineer who wrote the CPKAT software. Dr Weightman had responsibility for the overall conduct of the main study.

This thesis presents independent research commissioned by the National Institute for Health Research (NIHR) under a Clinical Research Doctoral Fellowship. The views expressed in this thesis are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health.

Other acknowledgements

I would like to thank my academic supervisors Professor Bipin Bhakta, Professor Martin Levesley, Professor Mark Mon-Williams and Dr Rory O'Connor for their consistent support, wisdom and guidance not only through this period of study but from the beginning of my research journey. I would like to place on record my recognition of the influence that Professor Bhakta has had on my life. Thanks to Professor Bhakta and Professor Levesley for taking the chance on me eight years ago and to Professor Mon-Williams and Dr O'Connor for their continued support in obtaining funding for me to continue a research career within their departments.

Thanks again to Professor Levesley for his support from the beginning of my research career and for his help, companionship and the use of his car during delivery and collection of the games devices to schools and children's homes.

Thanks to Dr Mike Clarke and Julie Cliffe for their valuable clinical supervision.

The Academic Department of Rehabilitation Medicine has benefitted greatly from professorial research and administrative assistants Karen Lawson and Vikki Wilkinson, and I would like to thank them for all their kindness, administrative and moral support to me over the years.

I would like to thank Dr Mike Clarke again, and also Dr Karen Pysden and Dr Clarke's medical team at the spasticity clinic at Leeds Teaching Hospitals NHS Trust, in particular Diane Doherty-Williams and Jeanette Wales.

Thanks to Jacqueline Lower, Dr Clarke's secretary, for her patience and for identifying each week those children that might prove eligible to participate in the study.

Thanks to each of the therapy teams that supported all aspects of the research described in this thesis. I would particularly like to mention therapy team managers in Leeds Community: physiotherapy manager Helen Dowden, and Jennifer McAnuff and Benita Powrie at the Leeds Community Children's Occupational Therapy Service. These inspirational team leaders have been very supportive of this research, which also included personal encouragement and moral support even when they were going through unsettling times themselves.

Almost 200 families have played a central part in this research. Thanks to each of them for their time and contribution, for their patience and hospitality, kindness and support.

Thanks to my family: my brother Simon, Uncle Robin and Aunt Judy who have been pillars of support all my life. Thanks to Suzie's parents Diane and Alf for everything that they are to Suzie and to Jack and to me, and for all the vital help they give us.

To Mike Horton and Dr Andy Weightman I owe an immense debt of gratitude. Dr Weightman has been a mentor since I first became involved in research, and has always been quick to lend an ear and to offer support, advice and a shoulder. Andy is an inspirational researcher, lecturer and friend and has played as influential a role as Professor Bhakta on supporting my progress and education as a researcher. Thanks for everything, Andy.

Mike Horton has been a roommate since the beginning, forced to share not only an office but also in all the adversities, reverses and despair of this thesis. Mike has never failed to pause and listen, and has provided much succour, humour, beer, tea, cakes, merriment and even shoes, ties and shirts. Mike has helped me immensely with Rasch-related research, but for everything, Mike, thank you very much.

During the course of the research described in this thesis, I have married and become a father. To Suzie and Jack, I owe the greatest thanks of all. I love you so much, I am so grateful for everything we have. You are what life is all about, and this thesis is a result of the support and focus you have given me. Thank you for everything.

Abstract

Background. Spastic cerebral palsy is a common cause of childhood activity limitation that restricts children's personal development. Botulinum toxin is a spasticity treatment that can improve upper limb activity limitation when combined with rehabilitation therapy. We investigated whether use of a computer-assisted arm rehabilitation (CAAR) device enhanced the benefits of botulinum toxin treatment of the upper limb of children with cerebral palsy.

Method. Fifteen children with cerebral palsy aged 5 – 12 years old undergoing botulinum toxin treatment for spasticity of the upper limb were randomly allocated into a CAAR group and a control group using minimisation, a procedure that balances groups for prognostic factors e.g. age and disability. Children used CAAR at home for 6 weeks. Assessments were carried out by blinded assessor at baseline, six and twelve weeks.

Primary outcome measure. ABILHAND-kids.

Secondary outcome measure. Canadian Occupational Performance Measure (COPM).

Results.

ABILHAND-kids. Activity limitation worsened following botulinum toxin treatment. An ANCOVA revealed that this was not significant (median scores, all participants: baseline, 0.8084; six weeks, 0.145; twelve weeks, 0.334; p=0.462) and that there was no difference between groups (p=0.699).

COPM. A Friedman's ANOVA revealed a statistically significant improvement that was clinically non-significant (baseline score, 4/10; six week score, 4.6/10; twelve week score, 4.6/10; p=0.031). A Kruskal-Wallis ANOVA revealed no difference in scores between groups at each time point.

CAAR use. Mean daily use, 7 minutes. Maximum use, 256 minutes (played over 24 days, mean daily use 10.667 minutes).

Sample size. This sample size was underpowered by 75%.

Conclusion. This study potentially supports evidence that botulinum treatment should be used only in combination with rehabilitation therapy but it was not adequately powered and a Type II error cannot be ruled out. The CAAR device did not engage the children enough to promote sufficient intensity and repetition of arm movements.

Table of Contents

List of Abbreviations

BiT	Bimanual Intensive Therapy
BMFM	Bilateral Fine Motor Function
CAAR	computer-assisted arm rehabilitation device
ChARM	Children's Arm Rehabilitation Measure
CIMT	Constraint Induced Movement Therapy
COPM	Canadian Outcome Performance Measure
CP	cerebral palsy
CPKAT	Cerebral Palsy Kinematic Assessment Tool
CRT	Charterhouse Rehabilitation Technologies
CTA	Clinical Trial Agreement
CTT	Classical Test Theory
DIF	Differential Item Functioning
ELBW	extremely low birth weight
GAS	Goal Attainment Scaling
GMFCS	Gross Motor Function Classification System
ICF-CY	International Classification for Functioning, Disability and Health for Children and Youth
IVH	Intraventricular Haemorrhage
LTF	Lost to Follow up
LTHT	Leeds Teaching Hospitals NHS Trust
MABC	Movement Assessment Battery for Children
MACS	Manual Ability Classification System
MAUULF	Melbourne Assessment of Unilateral Upper Limb Function
MHRA	Medicines and Healthcare products Regulatory Agency
mCIMT	modified Constraint Induced Movement Therapy
NIHR	National Institute for Health Research
PVL	periventricular leukomalacia
REC	Research Ethics Committee
R&D	Research and Development
SCPE	Surveillance of cerebral palsy in Europe
VLBW	Very Low Birth Weight
WHO	World Health Organisation

List of Figures

List of Tables

1 Cerebral palsy: background, classification, aetiology, epidemiology and upper limb rehabilitation

"Medicine is a science of uncertainty and an art of probability."

~ Sir William Osler, 1849 - 1919

1.1 Introduction

Cerebral palsy is one of the most common causes of disability in childhood. Primarily characterised by movement or postural impairment of varying severity and presentation, it is commonly but not always associated with cognitive and other impairments, and has a variety of causes. This heterogeneity has contributed to the lack of agreement on a definition and classification of the disorder by renowned scientists and clinicians for over 150 years.

The aim of this chapter is to provide a scientific basis for the thesis. It begins with an overview of cerebral palsy: its history and the attempts to define it; the causes of cerebral palsy; its epidemiology and the classification systems in common use to describe the presentation of cerebral palsy in children. This provides an essential background for describing the development and rationale for current approaches for the treatment and rehabilitation of the upper limb of children with cerebral palsy, and the theory underpinning the research study described in this thesis.

1.2 History

Today, cerebral palsy is recognised as an umbrella term for a motor disorder (Mutch et al., 1992) which may or may not be associated with other neurological impairments (Bax, 1964). It is a leading cause of disability in childhood (Reddihough and Collins, 2003). Cerebral palsy was first described as "cerebral paresis" in the mid-1850s by William Little, an orthopaedic surgeon specialising in tenotomy (Samilson, 1975, Stanley and Alberman, 1984), after he recognised that poliomyelitis was not always the cause of limb paresis presenting in many of the children attending his clinics. For this reason, the condition was first known as Little's Disease (Morris, 2007). It is recognised, however, that cerebral palsy is not a recent phenomenon. There are observations suggesting cerebral palsy in ancient texts (the Bible), and in ancient Egyptian monuments (McDonald and Chance, 1964). Little suggests that

cerebral palsy was the cause of the physical problems experienced by 15th century English monarch King Richard III, basing this diagnosis on the medical history (a difficult and premature birth) and descriptions published both in the historical records by Sir Thomas More (Longo and Ashwal, 1993) and by Shakespeare:

> *I that am curtailed of this fair proportion,*
>
> *Cheated of feature by dissembling Nature,*
>
> *Deform'd, unfinish'd, sent before my time*
>
> *Into this breathing world, scarce half made up,*
>
> *And that so lamely and unfashionable,*
>
> *That dogs bark at me as I halt by them*
>
> *(Richard III, Act 1, Scene 1(cited in Longo and Ashwal, 1993))*

Cerebral palsy had been attributed to evil spirits or a punishment by God (Aisen et al., 2011) but Little had noted a potential association between neurological lesions and the clinical presentation of children presenting with cerebral palsy, and their further association with child birth (Stanley and Alberman, 1984) and prematurity (O'Shea, 2008). Others associated the onset of the same clinical symptoms with illnesses such as mumps, measles or other infections (Stanley and Alberman, 1984). Following Little's pioneering work into the diagnosis and aetiology of cerebral palsy, it was investigated further by two other leading clinical scientists, Sigmund Freud and Sir William Osler (Samilson, 1975, Stanley and Alberman, 1984, Longo and Ashwal, 1993). Osler, a huge influence on the development of modern medical training, promoted the investigation of children with cerebral palsy as of special interest because he recognised that they were very responsive to medical care and "training" (Longo and Ashwal, 1993), the first mention of the impact of rehabilitation for children with cerebral palsy.

Classifying this "new" medical condition became an issue of both clinical presentation and aetiology. Sigmund Freud believed that it was impossible to classify on aetiological grounds and argued that only clinical findings were important (Morris, 2007), though he felt that with time and medical advances it would be possible to define cerebral palsy in terms of aetiology and underlying pathology as well (Stanley and Alberman, 1984). Freud also

proposed that prematurity was only an indication that something had already affected brain development (O'Shea, 2008). This was the beginning of a long-standing debate on the most accurate definition and classification of cerebral palsy. The debate continues today because of the wide range of causes, severity, movement impairments and other associated impairments that affect each child with cerebral palsy. The terms used to describe and classify children presenting with cerebral palsy have been in common use since they were coined by these pioneers in the examination and treatment of cerebral palsy, and only recently have investigators in the study of cerebral palsy recommended their replacement with modern terminology (see Table 1-1). However, their use is still common place within the literature and medical discussions.

Table 1-11. Terms in common use for describing children with cerebral palsy

Traditional (historical) use	Recommended new terminology[1, 2]	Description
Hemiplegia/ spastic hemiplegia	Unilateral cerebral palsy	Cerebral palsy with motor impairment affecting the limbs of one side of the body
Bilateral spastic hemiplegia	Bilateral cerebral palsy	Cerebral palsy with motor impairment affecting the limbs of both sides of the body
Spastic diplegia	Bilateral cerebral palsy affecting the lower limbs to a greater extent	Cerebral palsy with motor impairment affecting only or mostly the lower limbs
Paraplegia or quadriplegia	Bilateral cerebral palsy	Cerebral palsy with motor impairment affecting the limbs of both sides of the body

1 (Rosenbaum et al., 2007)

2 (Surveillance of Cerebral Palsy in Europe, 2000)

1.3 Definition and classification of cerebral palsy and its associated impairments

1.3.1 Introduction

The impact of Little, Osler and Freud in defining cerebral palsy as a separate group of disorders with a cerebral origin caused at or around birth and early childhood is well recognised (Longo and Ashwal, 1993, Morris, 2007, Samilson, 1975), and their descriptions

of the clinical presentation of the child with cerebral palsy have been used until recently (see Table 1-1) (Longo and Ashwal, 1993). But the debate to accurately define and classify cerebral palsy has continued into the 21st century, with contributions from leading experts on both sides of the Atlantic and progress impeded by difficulties such as the varying international interpretations of clinical terms, e.g. spasticity (Bax, 1964). Common definitions of these terms are given in Table 1-2.

Table 1-2. Definition of terms used to describe or classify children with cerebral palsy and their associated impairments

Term	Definition
Ataxic (movements)	Movements characterised by clumsiness, poor accuracy, poor stability (Paneth, 2008). Muscle tone tends to be low (Paneth, 2008)
Athetoid (movements)	Continuous, involuntary, dyskinetic movements characterised by writhing (Paneth, 2008, Rethlefsen et al., 2010).
Chorea (movements)	Dyskinetic movements which are quick and disjointed (Paneth, 2008), with bouts of continuous and indiscriminate involuntary movements (Rethlefsen et al., 2010).
Choreo-athetosis	Dyskinesia in which abnormal movements caused by spasticity dominate (Paneth, 2008) (a mixture of choreo-athetoid movements).
Dyskinesia	Involuntary limb movements which are exaggerated with voluntary attempted movements (Paneth, 2008).
Dystonia	Dyskinesia in which abnormal postures caused by spasticity dominate (Paneth, 2008).
Spasticity	A pathological increase in the muscle stretch reflexes that is velocity-dependent (Voerman et al., 2005) i.e. passive movement of a joint meets with resistance that is proportional to the speed of movement (Rethlefsen et al., 2010). Characterised by sustained muscle contractions and exaggerated tendon jerks (Thompson et al., 2005).

1.3.2 Definition of cerebral palsy

To date, and despite the international collaborations and workshops to address this subject (Blair et al., 2007, Cans et al., 2007, Rosenbaum et al., 2007), there is still no globally-accepted definition for cerebral palsy. This lack of agreement reflects the heterogeneity of the disorder. Two international authorities on cerebral palsy agree on four essential

components for an accurate definition (Blair et al., 2007, Surveillance of Cerebral Palsy in Europe, 2000):

- cerebral palsy is an “umbrella term” for a group of disorders;
- it is permanent but not unchanging;
- it is a movement/postural and motor function disorder;
- it is due to a non-progressive interference/lesion/abnormality in the developing/immature brain.

However, Blair et al (2007) do not think that these criteria are detailed enough, and describe additional specific factors that are the basis for many of the disagreements. These include a defined minimum level of movement/postural disability; lower and upper age limits for when cerebral palsy can develop and be diagnosed; and a definition of the age at which the brain is said to be ‘mature’ (Blair et al., 2007). Other leading commentators highlight the failure, still, to include aetiology, and its exact timing (Alberman and Mutch, 2007).

Some of the more commonly-accepted or more robustly developed definitions of cerebral palsy are given in Table 1-3.

Table 1-3. Definitions of cerebral palsy.

"A disorder of movement and posture due to a defect or lesion of the immature brain." (Bax, 1964)
" an umbrella term covering a group of non-progressive, but often changing, motor impairment syndromes secondary to lesions or anomalies of the brain arising in the early stages of its development." (Mutch et al., 1992)
"Cerebral palsy describes a group of permanent disorders of the development of movement and posture, causing activity limitation, that is attributed to non-progressive disturbances that occurred in the developing foetal or infant brain. The motor disorders of cerebral palsy are often accompanied by disturbances of sensation, perception, cognition, communication, and behaviour, by epilepsy, and by secondary musculoskeletal problems." (Rosenbaum et al., 2007)
"Cerebral palsy is a group of permanent, but not unchanging, disorders of movement and/or posture and of motor function, which are due to a non-progressive interference, lesion, or abnormality of the developing immature brain." (Cans et al., 2007)

None of these is defendable against the criticisms outlined above by Blair et al (2007) or Alberman and Mutch (2007), but they are the result of between 100 and 150 years of pooled international resources from some of the most distinguished clinicians and scientists in history, suggesting that a definition of cerebral palsy that is acceptable to everyone involved in treating children with cerebral palsy is still unlikely.

1.3.3 Classification of cerebral palsy

Classification "is the process of grouping data, persons, or objects into classes according to common characteristics" (Eliasson et al., 2006, p. 549). Classifying cerebral palsy serves a number of useful purposes.

1. It provides essential information about the presentation of the patient for diagnostics, investigations and treatment (McLaughlin, 2007, Rosenbaum et al., 2007).

2. It guides current and future referrals to health services and informs health service managers and clinicians of service requirements (McLaughlin, 2007, Rosenbaum et al., 2007).
3. It allows accurate evaluation of change at different time points (Rosenbaum et al., 2007).
4. It supports robust design of research study protocols (McLaughlin, 2007).

Classification of cerebral palsy, like the definition of cerebral palsy, has also proved difficult because of the wide range of components on which classification could be based. For example, throughout the 20th century and into the 21st century, contributors have argued for inclusion of the following components (Colver, 2007, Graham, 2007, Morris, 2007, Perlstein, 1952):

- neurological;
- cognitive impairment;
- motor function (lower limb);
- upper limb function (manual ability);
- anatomical location of the brain lesion;
- clinical symptoms (e.g. spasticity, athetosis, rigidity etc.);
- quantification of spasticity;
- topography (anatomic distribution);
- severity of muscle tone;
- the severity of involvement;
- aetiology;
- quality of life;
- participation in life situations;
- therapeutic requirement or treatment protocol.

As an example of a classification based on topography and clinical presentation, Mutch et al (1992) proposed a modification of a Swedish classification of cerebral palsy, shown in Table 1-4, that continued to use terms established in the previous century.

Table 1-4. Modified Swedish classification of cerebral palsy based on topography (Mutch et al., 1992)

Spastic cerebral palsy	Hemiplegia
	Tetraplegia
	Diplegia
Ataxic cerebral palsy	Diplegia
	Congenital
Dyskinetic cerebral palsy	Mainly choreoathetotic
	Mainly dystonic

Mutch et al (1992) further emphasised the requirement for an aetiologically-based classification but echoed Freud's views from over a century before that this was unlikely without the use of advanced medical technology. Mutch et al's (1992) taxonomy also serves to illustrate other issues involved in classifying cerebral palsy: the terms included within it are the subject of further debate. The report of the international workshop on the Definition and Classification of Cerebral Palsy included strong recommendations that the classifications of diplegia, hemiplegia, tetraplegia and quadriplegia are replaced by the terms bilateral and unilateral motor involvement (Rosenbaum et al., 2007), as described and already in use by Surveillance of Cerebral Palsy in Europe (Surveillance of Cerebral Palsy in Europe, 2000). Dammann and Kuban (2007) argue against dropping the older classifications while Graham (2007) supports adoption of the new terms, stating that he has found a large variation in classification of bilateral types using the older classifications.

Despite formulating one of the most widely accepted definitions of cerebral palsy (see Table 1-3), Bax (1964) admitted that his informal collaboration had found it difficult to agree about any classification of cerebral palsy. The international workshop (Rosenbaum et al., 2007) did not fully resolve this issue either, but did issue recommendations that any classification of cerebral palsy should incorporate the categories shown in Table 1-5.

Table 1-5. Components of cerebral palsy classification recommended by the international workshop to define and classify cerebral palsy (Rosenbaum et al., 2007)

Motor abnormalities	
Nature and typology of the motor disorder	The observed tonal and movement abnormalities
Functional motor abilities	Motor function limitations (gross motor function and manual ability)
Accompanying impairments	
Sensation, perception, cognition, communication or musculoskeletal problems.	
Anatomical and neuro-imaging findings	
Anatomic distribution (topography)	The parts of the body affected e.g. lower limbs
Neuro-imaging findings	Report of MRI or CT imaging
Causation and timing	
What is the cause and when did it occur	

1.3.3.1 Motor skills classifications

Two functional motor ability classification systems (for manual ability and gross motor function) are gaining recognition and support (Morris and Bartlett, 2004, Rosenbaum et al., 2007, Surveillance of Cerebral Palsy in Europe, 2000): the Manual Ability Classification System (MACS; Eliasson et al., 2006) and the Gross Motor Function Classification System (GMFCS; Morris and Bartlett, 2004) respectively, though the Surveillance of Cerebral Palsy in Europe (SCPE) endorses the Bilateral Fine Motor Function (BFMF; Beckung and Hagberg, 2002) over the MACS for classifying manual ability. The advantages of the BFMF are that it accounts more accurately for differences of unilateral impairment over bilateral impairment and it can be more easily scored from medical notes (Cans et al., 2007). SCPE also state reservations about these classifications, in that intellectual impairment has an effect on manual and motor ability and this influences the scores (Cans et al., 2007). The levels which define these classification systems are given in Table 1-6.

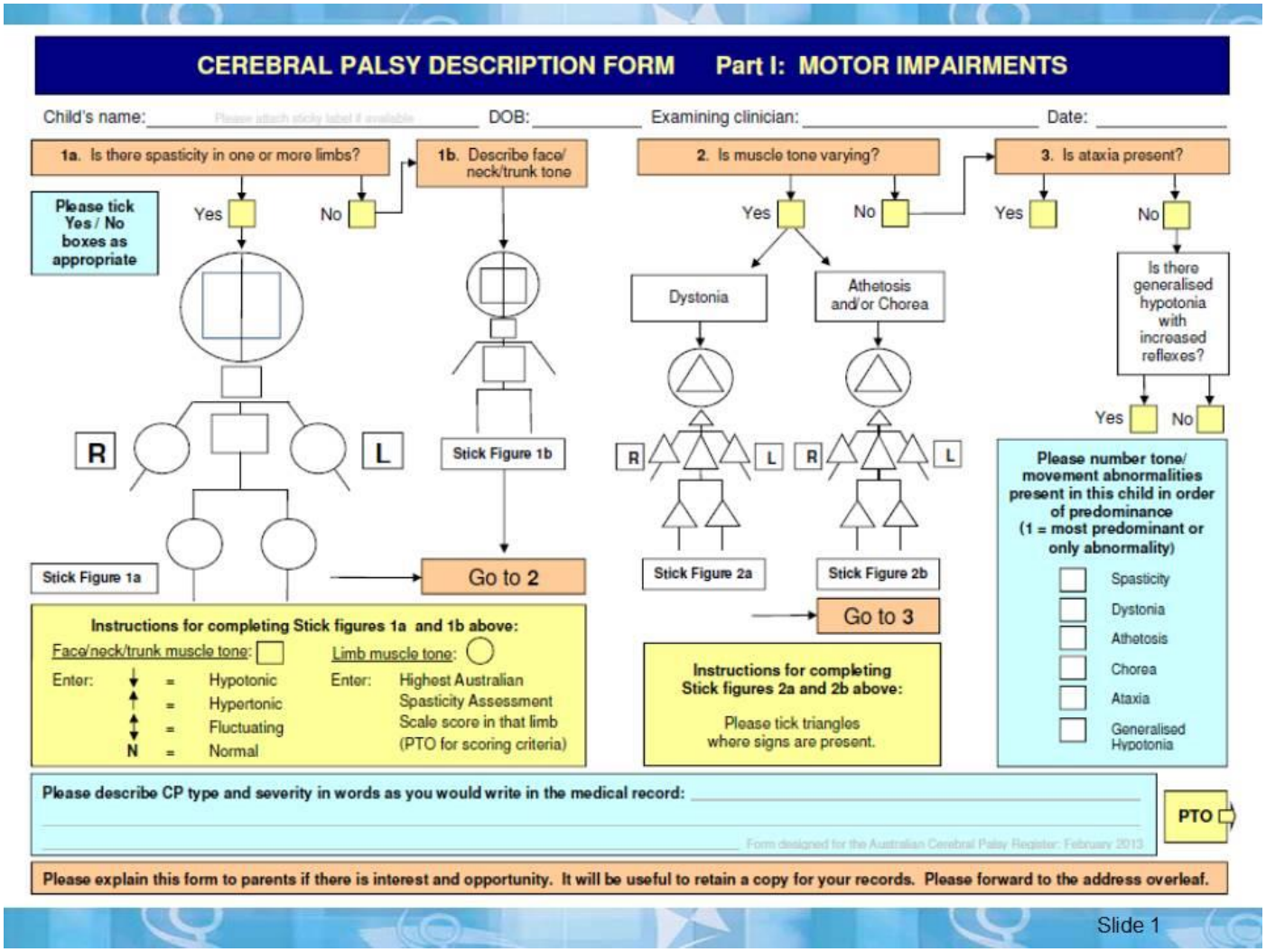
Table 1-6. Commonly-used classification systems for manual and gross motor function in cerebral palsy.

Classification/ Levels	Manual Ability Classification System (MACS) (Eliasson et al., 2006)	Gross Motor Function Classification System (GMFCS) (Palisano et al., 1997)	Bilateral Fine Motor Function (BFMF) (Beckung and Hagberg, 2002)
Level I	Handles objects easily and successfully.	Walks without Limitations	One hand: manipulates without restrictions. The other hand: manipulates without restrictions or limitations in more advanced fine motor skills.
Level II	Handles most objects but with somewhat reduced quality and/or speed of achievement.	Walks with Limitations	(a) One hand: manipulates without restrictions. The other hand: only ability to grasp or hold (b) Both hands: limitations in more advanced fine motor skills
Level III	Handles objects with difficulty; needs help to prepare and/or modify activities.	Walks Using a Hand-Held Mobility Device	(a) One hand: manipulates without restrictions. The other hand no functional ability (b) One hand: limitations in more advanced fine motor skills. The other hand: only ability to grasp or worse
Level IV	Handles a limited selection of easily managed objects in adapted situations.	Self-Mobility with Limitations; May Use Powered Mobility	(a) Both hands: only ability to grasp (b) One hand: only ability to grasp. The other hand: only ability to hold or worse
Level V	Does not handle objects and has severely limited ability to perform even simple actions.	Transported in a Manual Wheelchair	Both hands: only ability to hold or worse

Other impairments experienced by children with cerebral palsy (e.g. for communication, feeding and swallowing disorders) are the subject of other classification systems (Sršen, 2012).

Himmelman et al (2006) argue that classification of cerebral palsy should be based on the topography, as listed in Table 1-4, and motor function using the GMFCS, as this combination gives a good general impression of the overall clinical presentation (of impairment). Their findings support the earlier work of co-authors (Beckung and Hagberg, 2002) who noted moderate associations between gross motor performance, manual ability and cognitive impairment (Beckung and Hagberg, 2002). This argument is supported independently by a collaboration of experts in Australia involved with the Australian Cerebral Palsy Register (ACPR; Blair et al., 2007)) but in order to facilitate greater communication and data collection for classification, they have developed a form (see Figure 1-1, used with permission).

Figure 1-1. Australian Cerebral Palsy Register data collection form.



Part 2 of the Australian Cerebral Palsy Register data collection form requires a description of the child's functional ability using the GMFCS and the MACS, and their associated impairments.

Global agreement on classification of cerebral palsy therefore includes a description of the child's motor impairments, topography and their functional limitations. A description of accompanying impairments may also be included but imaging investigations and details of the cause may be difficult to obtain in many countries e.g. those with limited medical resources.

1.3.3.2 International Classification for Functioning, Disability and Health for Children and Youth

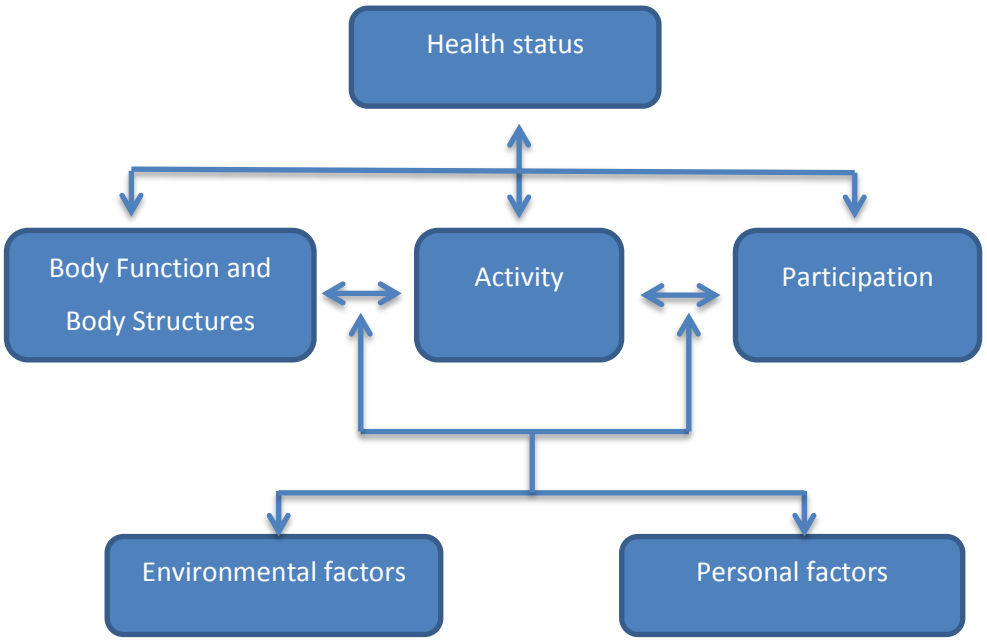
Attempts to classify cerebral palsy have focused on the clinical presentation and causes. Although there are still arguments for the inclusion of aetiology, the consensus appears to be that a description of impairment, functional limitations and the topography are essential. This is unsurprising as a basis for classification of cerebral palsy, because they stem from the historical basis for the treatment of disease and disability: the medical model. In the medical model the cause of disease and the disabling consequences of the cause are identified and targeted, thus effecting an improvement in function and health of the patient (Majnemer and Mazer, 2004). Although this model was useful in some circumstances of ill health e.g. for the treatment of infection with antibiotics (Rosenbaum and Stewart, 2004), it does not meet with universal approval. Firstly, it offers only an inflexible, constrained single direction pathway that focuses attention on the negative consequences for the person (Rosenbaum and Stewart, 2004, Mittrach et al., 2008); secondly, it is limited in its scope (Hemmingsson and Jonsson, 2005). The medical model places the medical team as central to addressing the disease and disabling consequences.

This situation was not acceptable to people with disabilities. The medical model placed their impairment and disability as a problem that was entirely theirs to deal with, albeit supported by a medical team. The effect of this model was to isolate disabled people from society. People with disabilities felt discriminated against, because they were restricted from participating in many life situations: schools; universities; jobs; leisure and cultural

activities. However, most people suffer from an impairment of some type in their lifetimes. This might be as a result of aging in previously healthy people e.g. deteriorating eyesight and hearing, or increasing mobility impairment and infirmity. Temporary disability is also common e.g. reliance on a wheelchair after a sporting or workplace injury that, while short-lived, causes the sort of participation restriction that is a permanent aspect of disabled people's daily lives. Disability was, disability groups pointed out, a problem for society as a whole. A social model of disability was proposed, in which problems associated with having a disability were caused by society (Oliver et al., 2012). In this way, stairs to gain entry to a building are an obstacle not because of mobility impairment but because of society's focus on the ability of the majority, which discriminates against disabled people. The social model proposed therefore that attitudes, legislation and society itself needed to address disease and disabling consequences to allow inclusion and participation of all people who suffer from any form of disability, whether permanent or temporary.

During the 1990s, the World Health Organisation (WHO) developed a new model (Rosenbaum and Stewart, 2004): the International Classification for Functioning, Disability and Health for Children and Youth (ICF-CY) which was ratified by the World Medical Association in 2001 (World Health Organisation, 2001). The ICF-CY relates to all people, not just those with ill health (Colver, 2007, Palisano et al., 2004). It features domains that relate to people's health, well-being and physical, social and economic activity. The ICF-CY model, illustrated in Figure 1-2 below, also emphasises that there is an interaction of these four domains, underlining the contrast with the linear dependency on treatment for the disease implied by the medical model (Rosenbaum and Stewart, 2004).

Figure 1-1. The International Classification of Functioning, Disability and Health



The ICF (World Health Organisation, 2001) defines Body Functions as the physiological functions of body systems; Body Structures are the anatomical parts of the body (e.g. organs, limbs); activity is the execution of a task or action by a child at the individual level; participation is involvement in life situations, from the viewpoint of society (Stucki and Sigl, 2003). The ICF further defines activity limitation as any difficulty experienced by the child in carrying out a task or action, and participation limitation is any difficulty experienced by a child in participating in a life situation, measured against normal peer expectations (Stucki and Sigl, 2003). Contextual factors include Environmental and Personal factors. Environmental factors include the physical environment – the natural and man-made features - and other people, society's values, health services, and legislation. Personal factors relate to each individual child e.g. their age, gender and socioeconomic background. The ICF-CY recognises disability as defined by the UN Convention on the Rights of Persons with Disabilities as "those who have long-term physical, mental, intellectual or sensory impairments which in interaction with various barriers may hinder their full and effective participation in society on an equal basis with others (United Nations, 2006). This definition and the ICF-CY views the child's health and wellbeing as a relationship between:

- his or her medical status;
- the domains of Body Structures and Body Function;
- Activity and Participation;
- environmental and personal factors

It follows then that improvement in a child's health and wellbeing or a reduction in activity limitation or participation restriction can be achieved through management of environmental and personal factors (Colver, 2007, Stucki and Sigl, 2003).

The ICF-CY is therefore a middle ground between the social model and the medical model. It is a classification system that changes the focus from the disability to the contextual factors (Colver, 2007, Hemmingsson and Jonsson, 2005, Palisano et al., 2004). A commonly-used illustration of this model is the re-arrangement of tutoring in an upstairs classroom to one on the ground floor to accommodate a pupil with cerebral palsy who has mobility problems (Palisano et al., 2004). In such situations, participation in life situations relevant to the child's wellbeing are successfully enhanced without recourse to medical treatment but by addressing environmental obstacles (Rosenbaum and Stewart, 2004). This change of approach recognises the necessity to adapt and develop new outcome measures that evaluate change in quality of life and participation restriction following intervention (Majnemer and Mazer, 2004) but it also potentially changes the focus of classification of the child with cerebral palsy from impairment and clinical description to activity restriction and participation limitation, and the impact of environmental and personal factors. This has the potential to support the classifying of all health conditions with respect to the treatments and the eventual targeted outcomes, as argued by Damiano (2007). An example of this is given by Preston et al. (2011) in the goals and outcomes of a spasticity clinic for children with cerebral palsy (see Table 1-7).

Table 1-7. Example of use of ICF-CY to categorise goals of treatment.

Goals of treatment (ICF-CY domain and category)	Clinically assessed outcomes of treatment (ICF-CY domain and category)
Improve hand function (d440 Fine hand use)	Grips objects better (d4401 Grasping)

ICF-CY: International Classification for Functioning, Disability and Health for Children and Youth

1.3.4 Summary and conclusion: definition and classification of cerebral palsy

The global debate about an acceptable definition and classification of cerebral palsy has continued for over 150 years, and no agreement appears likely in the second decade of the 21st century. Arguments about what to include in both the definition and classification

have been driven in part by the heterogeneity of the condition, encompassing as it does the possibility of a wide variety of associated impairments as well as the primary motor disorder. Classification should always include:

- the topography and nature of movement disorder;
- the motor impairment and severity and any associated morbidity;
- the inclusion of neuroimaging.

Inclusion of the cause and its timing is recommended but is part of the ongoing debate; however, these can not be substantiated or investigated in areas where medical technology or detailed notes are unavailable. The WHO's International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY) is changing the approach of society and health care towards chronic disease and disability. The ICF-CY focuses on factors such as personal and environmental factors, and emphasises health rather than disability and impairment. This places the emphasis on changing social attitudes and physical barriers to enable greater involvement in society and life situations for all people with long term health issues, and might yet influence how health care classifies the child with disability.

Within this thesis, it will be necessary to give a description of the children with cerebral palsy that have supported the research. Their characteristics are essential for sampling purposes and for obtaining a representative sample for the various aims of this work. The children will therefore be classified using topography and severity, gross motor and manual ability limitations and any associated impairments. Outcomes will be captured using activity limitation measures.

1.4 Aetiology

1.4.1 Introduction

Although a single, universally-acceptable definition of cerebral palsy has still not been achieved, there is consensus that it must include the cause (see Table 1-3). This section explores the risk factors and the aetiology of cerebral palsy and relates findings to the different clinical presentations. Unsurprisingly, when one considers the heterogeneous nature of cerebral palsy, the areas of disturbance to the brain tissue are widespread and inconsistent (Stanley and Alberman, 1984). For many years this was thought to be due to hypoxia, but subsequent studies have identified other risk factors and potential causes.

1.4.2 The earliest hypothesis for cerebral palsy – hypoxia caused by child birth

1.4.2.1 Background

Neural tissue is oxygen and nutrient-hungry. In adults the brain makes up 2% of the total body mass yet uses 20% of the body's total oxygen requirement (Zauner and Muizelaar, 1997). It also receives a sixth of cardiac output and a quarter of the body's glucose (Zauner and Muizelaar, 1997). The foetal brain makes up 12% of total body mass (Stanley and Alberman, 1984) therefore it seems reasonable to suggest that foetal cerebral tissue is particularly vulnerable in the event of hypoxia (Stanley and Alberman, 1984).

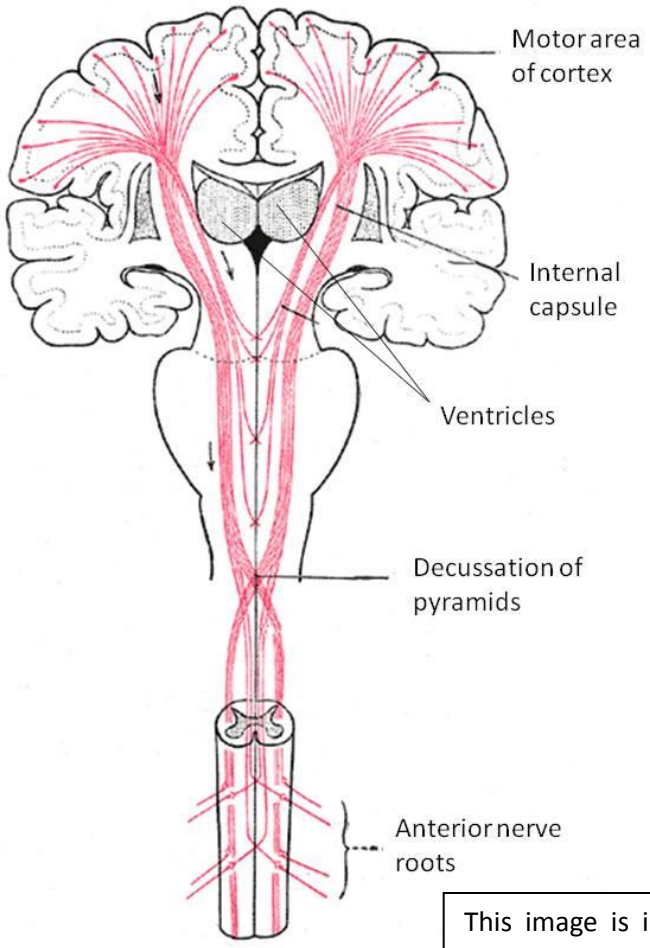
1.4.2.2 Causes of hypoxia in cerebral tissue: childbirth and prematurity

Problems with childbirth were thought to be the cause of cerebral palsy in the majority of children from the time of William Little until recently (Reddihough and Collins, 2003). Cerebral tissue hypoxia results from haemorrhage caused by damage to tissues during birth trauma, or from ischemia (Stanley and Alberman, 1984); both haemorrhage and ischemia have the same effect, that of starving tissue of oxygen and nutrients (Stanley and Alberman, 1984). Additionally, it is likely that immature brain tissue and blood vessels are particularly weak and prone to trauma (Aisen et al., 2011). Cerebral hypoxia is likely to cause an increase in cerebral blood flow as a physiological response to low oxygen levels in the tissue. However, the increase in cerebral blood flow, if continued for an extended period, can result in cerebral oedema that is itself a cause of hypoxia by reducing brain perfusion (Stanley and Alberman, 1984). Hypoxia also causes reduced cardiac output (Miller and Clark, 1998); the resultant drop in blood pressure worsens the delivery of oxygen to the cerebral tissue at the peripheries of the circulatory system (known as the watershed) (Stanley and Alberman, 1984). These peripheries include white matter close to the cortex, the periventricular areas and the basal ganglia (Stanley and Alberman, 1984) which cause spastic bilateral cerebral palsy, bilateral cerebral palsy affecting the legs more than the arms and choreoathetoid cerebral palsy respectively (Miller and Clark, 1998, Koman et al., 2004).

Bleeding into the ventricles (intraventricular haemorrhage) or into the brain tissue around the ventricles (subependymal haemorrhage) is common in almost half of all premature newborns (defined as birth between 20 and 37 weeks gestation; Wisanskoonwong et al., 2011) of less than 1500 grams (Stanley and Alberman, 1984), causing disturbances to the periventricular white matter (Aisen et al., 2011). The developing brain in the

periventricular area is also susceptible to ischaemic damage (Stanley and Alberman, 1984). Lesions observed in this area are termed periventricular leukomalacia. Both intraventricular haemorrhage and periventricular leukomalacia (but more so periventricular leukomalacia; Miller and Clark, 1998) are each strongly associated with cerebral palsy (Jacobsson and Hagberg, 2004). They indicate destruction of white matter and impede development of the corticospinal pathways (Stanley and Alberman, 1984) and are the most common abnormal characteristic found in children with cerebral palsy (Robinson et al., 2009, Shevell et al., 2003, Krageloh-Mann and Horber, 2007). The proximity of these corticospinal tracts to the ventricles is shown in Figure 1-3. Note how the tracts closest to the ventricles are associated with the motor neurone pathways supplying the lower limbs so that damage restricted to this region results in impairment of lower limb motor function. Note also the descending tracts. These cross over at the decussation of the pyramids to supply the contralateral side of the body, illustrating how cerebral injury results in impairment to the contralateral limbs. At birth, there are ipsilateral neural connections which reduce over the first two years as motor skills develop (Kirton, 2013), but about 10% of these nerve fibres remain to supply the ipsilateral side (FitzGerald et al., 2012). This has implications for motor function and motor rehabilitation (Andersen et al., 2013), which will be discussed later in section 1.6.4.

Figure 1-3. Diagram from Gray's Anatomy showing the position of the ventricles and the proximity to them of the descending corticospinal tracts.



1.4.2.3 Doubts about prematurity and hypoxia as causes

While hypoxic events have clear effects on cerebral tissue and there is an implication of preterm birth as a predictive factor for cerebral palsy (O'Shea, 2008, Blair and Watson, 2006), they account for less than 50% of cases (Koman et al., 2004). Nelson and Grether (1999) point out that cerebral palsy is more common in full term infants than premature children. With the advances in obstetrics and neonatal medicine for both preterm and term babies, there has been a decrease in stillbirth and birthing mortality but no decrease in the incidence of cerebral palsy (Reddihough and Collins, 2003, Jacobsson and Hagberg,

2004). Reddihough and Collins (2003) assert that cerebral palsy is caused by perinatal asphyxia in only 8% of cases, and suggest that the cause of cerebral palsy is not premature child birth but that prematurity has itself been precipitated by the cause of cerebral palsy. Jacobsson and Hagberg (2004) further quantify this as 70 – 80% of cases of cerebral palsy being caused by prenatal factors. There is, therefore, growing consideration given to proposals that other factors play a greater role in the development of cerebral palsy than birth asphyxia or prematurity itself, and that birth asphyxia and prematurity are themselves precipitated by the causative factors.

1.4.3 Other risk factors

Over 150 years ago, William Little first proposed that a difficult birth was a leading cause of cerebral palsy (O'Shea, 2008), and there are suggestions that a preterm birth less than 28 weeks gestation increases absolute risk of cerebral palsy by a factor of 100 (O'Shea, 2008). However, most children with cerebral palsy are not premature (Nelson and Grether, 1999) and improved obstetric care, advances of modern medicine in neonatology, monitoring heart rate and attempts to prevent birth asphyxia (Koman et al., 2004, O'Shea, 2008, Blair and Watson, 2006) have not improved the prevalence of cerebral palsy. There remains a clear risk of cerebral damage in the event of hypoxia (whatever the reason), but cerebral palsy may have a prenatal cause in 70 – 80% of cases (Jacobsson and Hagberg, 2004, Goldenberg et al., 2008). Table 1-8 shows typical risk factors for cerebral palsy that have been identified by a number of studies. Most authors on the subject have also identified a birth weight of less than 1,500 grams as a strong risk factor, e.g. Grether et al (1996) who find the risk a hundred times greater than in infants born over 3,000 grams.

Table 1-8. Risk factors associated with cerebral palsy

(Nelson and Ellenberg, 1986)	(Koman et al., 2004)	(Blair and Watson, 2006)	(Grether et al., 1996)	(Jacobsson and Hagberg, 2004)
Before pregnancy				
	Multiple births	Multiple births		Multiple births
	Foetal infection	Infection of the foetal membranes (white infants only)		Intrauterine infection (e.g. rubella, CMV), and inflammation
Motor disorder in older sibling				Cerebral palsy in older sibling
Hyper-thyroidism				Iodine deficiency
Maternal seizures				Low maternal age (< 20 years)
Previous unsuccessful pregnancies (more than 2)				Advanced maternal age (> 35 years)
Mother's learning disability				Maternal learning disability, epilepsy, diabetes

				Low socio-economic status
During pregnancy				
Severe proteinuria				
Bleeding in the third trimester				
Thyroid and oestrogen use				Maternal thyroid deficiencies
Asymptomatic heart disease				
Incompetent cervix				
Rubella				
During labour and delivery				
Gestational age of less than 33 weeks	Second stage of labour lasting more than 4 hours	Gestational age (rates of 27 weeks or less are double those from 28 – 32 weeks)	Preterm labour and rupture of membranes within 2 hours of hospital admission	Low gestational age
Low foetal heart rate (less than 61 beats per minute)	Vaginal bleeding		Vaginal bleeding at time of admission	Male gender
Breech presentation	Foetal anoxia		Place of birth (lack of specialised care)	Ruptured placenta

Chorio-amnionitis	Chorio-amnionitis	Chorio-amnionitis	Fever , chorio-amnionitis
Small placenta (weight <325 grams) or placental complications			Use of antibiotics
Postnatal period (delivery room)			
Birth weight <2000 grams		Birth weight <1500 grams	
Time to cry > 5 minutes			
Asymmetrical Moro's reflex			
Caucasian			
Micro-encephaly			
Post-delivery room period			
Neonatal seizures	Hyper-bilirubinaemia	Hyperoxia	Neonatal seizures
Major non-CNS malformations		Hypocapnia	
Maternal infection	Maternal infection		
Antibiotics given without infection			

Blair and Watson (2006) suggest that the most obvious risk factors are also those most likely to identify a *pathway* for causation, acknowledging that the risk factor itself might not be the cause. One such pathway involves maternal infection, for which there is a correlation with periventricular leukomalacia (Jacobsson and Hagberg, 2004) and which may cause between 20 and 40% of preterm births (Smith et al., 2009). Odding (2006) suggests that seasonal variations of disease-causing organisms may explain the increase in summer births of children who develop cerebral palsy, which adds weight to the idea of maternal infection as a cause for cerebral palsy. The pathway proposes that maternal infection precipitates raised foetal inflammatory agents (cytokines) and these are responsible for the cerebral (white matter) damage (Dammann and Leviton, 1998) that presents as periventricular leukomalacia. Use of anti-cytokines has implications for the normal brain development, however, because they are known to be involved in immune response and restricting cerebral tissue injury (Nelson and Grether, 1999). A systematic review by Smith et al (2009) which examines interventions aiming to prevent prematurity finds that prophylactic maternal antibiotics do not prevent preterm labour in mothers whose membranes remain intact; the review does, however, provide poor quality evidence that prophylactic maternal antibiotics do prevent inflammation of those membranes and reduce the number of neonatal adverse events. This is supported by Wisanskoonwong et al (2011), who note some benefit of antibiotics in a sub-group of women with a history of premature birth and a very low weight (less than 50 kg).

Wisanskoonwong et al (2011) points out that many risk factors are socio-economic e.g. poverty, education and exercise and could be tackled without medical intervention. Other known causes are already preventable e.g. inter-family marriages, iodine deficiency, rhesus isoimmunisation (Blair and Watson, 2006). Some therapies that increase survival in premature babies are themselves associated with cerebral palsy e.g. steroids to prevent chronic lung disease and neonatal ventilation (Blair and Watson, 2006), raising further potential areas for investigation.

1.4.4 Summary of aetiology

Recognition of the lesser role played by birth asphyxia and prematurity has renewed interest into potential causes of cerebral palsy. The most promising of these are the treatment or prevention of infection during pregnancy, investigation and improvement of interventions in neonatal intensive care units, and prevention of haemorrhage. There is still no definitive cause or causes of cerebral palsy but if targeting risk factors can minimise

the damage to cerebral tissue, benefits of rehabilitation to reduce activity limitation are likely to improve, and participation in life situations could increase favourably.

1.5 Epidemiology, prognosis and demographic characteristics

1.5.1 Introduction

Cerebral palsy is the most common cause of disability in children (Reddihough and Collins, 2003). A number of studies have investigated the prevalence of cerebral palsy and the implications for body function and structures, activity and participation. Knowing the probability and extent of the child's activity limitation, participation restriction or associated impairments helps to plan for the level of services and support required, and there is evidence that these prognostic indicators can help parents to cope (Novak et al., 2012).

The purpose of this section is to describe the prevalence of cerebral palsy in the developed world and the likelihood of types and severity of activity limitation, participation restriction and associated impairments that characterise cerebral palsy.

1.5.2 Prevalence

In 2002, a network of registers known as the Surveillance of Cerebral Palsy in Europe (SCPE; Surveillance of Cerebral Palsy in Europe, 2000) reported the birth statistics from 13 areas across Europe (Johnson, 2002) Their report focused on the prevalence and characteristics of children with cerebral palsy several of which were UK-based. SCPE developed an extensive dataset on 6,502 children with cerebral palsy.

This report demonstrated an overall prevalence of 2.1 children with cerebral palsy per 1000 live births (95% CI: 2.02 to 2.14), though some sites fell outside the confidence interval (Johnson, 2002). This variation might be due to differences of socio-economic status between areas, a recognised risk factor for cerebral palsy (Jacobsson and Hagberg, 2004, Wisanskoonwong et al., 2011) which increases the prevalence in areas of the UK to 3.33 per 1000 births (Odding et al., 2006). The ratio of boys with cerebral palsy to girls is 1.33, supporting the finding by Jacobsson and Hagberg (2004) that being a male carries a greater risk. Of 3,434 children for whom data was included, 73% of children were less than 1,500 grams at birth, again confirming this as a strong risk factor – only 1.2% of children above 2,500 grams developed the disorder (Johnson, 2002). Severe cerebral palsy (defined as an IQ less than 50, and non-ambulant) across Europe was 0.43 per 1,000 live birth (95% CI 0.4 to 0.46; Johnson, 2002).

1.5.3 Topography

The literature shows a number of population-based studies from various countries that identify the proportion of subtypes of cerebral palsy classified by anatomical distribution, but these vary considerably in their findings. The largest and most recent study in Europe that includes a number of UK sites finds that bilateral spastic cerebral palsy (at 55%) is the most common subtype (but it does not identify the proportion that presents with only lower limb involvement) and 29% had unilateral impairment (Johnson, 2002).

An extensive literature search by Odding et al (2006) supports this but Odding et al. (2006) break the proportions down further. To summarise their findings: dyskinetic cerebral palsy affects 12 – 14% of children, ataxic cerebral palsy affects 4 – 13% of children and spastic cerebral palsy affects up to 91% of children with cerebral palsy (Odding et al., 2006). Of the 91% of children with spastic cerebral palsy, only the lower limbs are involved in 13% to 25% of children (Odding et al., 2006). Odding et al. (2006) describe unilateral impairment (including one upper limb) presenting in up to 40% of children, with up to 43% of children having both upper limbs affected (bilateral cerebral palsy involving all four limbs). In total, therefore, upper limb impairment affects up to 83% of children with spastic cerebral palsy.

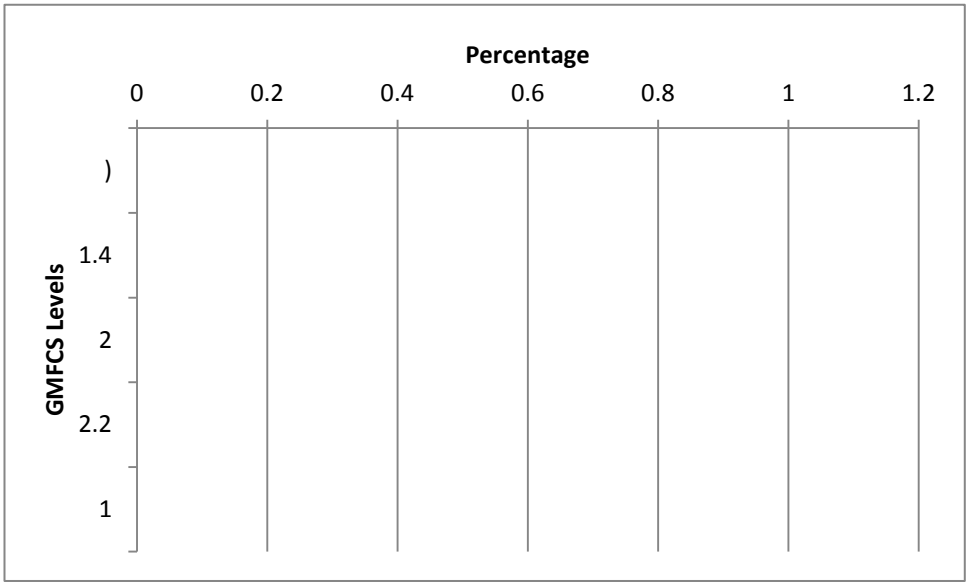
1.5.4 Activity limitation and participation restriction: motor and associated impairments

1.5.4.1 Motor impairments

Every child with cerebral palsy has a motor impairment. The discussion about classification of cerebral palsy describes some commonly-used fine and gross motor classification systems: the GMFCS for mobility, the MACS and the BFMF for manual ability (see section 1.3.3). In possibly the largest review of cerebral palsy undertaken, Novak et al (2012) found that, in almost 3,000 cases, over half the children had independent mobility, with a total of 36% walking with no or minimal motor impairment (GMFCS Level I). About half of this figure (17%) were unable to walk at all (Novak et al., 2012). Beckung and Hagberg (2002) identified a strong correlation (r = 0.74, $p<0.0001$) between the GMFCS and the BFMF (see Table 1-6), suggesting that the severity of mobility limitation is commonly associated with the severity of manual ability limitation.

Figure 1-4 shows the proportion of GMFCS levels in children with cerebral palsy (Novak et al., 2012).

Figure 1-4. Proportion of GMFCS levels in children with cerebral palsy.



Carnahan et al (2007) found a less strong correlation between MACS and GMFCS (kappa value 0.35, 95% confidence interval 0.27– 0.41) but noted that it depended on the cerebral palsy topography (limb distribution). Unilaterally-impaired children were more likely to have severe manual impairment compared to mobility impairment, and children with bilateral lower limb impairment, clearly, had greater mobility impairment on the GMFCS than manual ability impairment. However, there is a strong likelihood that those children who have severely limited mobility or use walking aids (GMFCS Levels III and IV) will lose the ability to walk in adolescence (Novak et al., 2012).

1.5.4.2 Associated impairments

The definition of cerebral palsy given by Rosenbaum et al (2007; see Table 3 7) is the only one that gives a description of the impairments commonly associated with cerebral palsy. A number of studies have attempted to quantify these impairments in the cerebral palsy population (e.g. Beckung and Hagberg, 2002, Himmelmann et al., 2006, Novak et al., 2012, Odding et al., 2006), and this section will summarise the proportions of these impairments presenting in children with cerebral palsy.

1.5.4.2.1 Learning disability

Twenty-three to 44% of children with cerebral palsy have a cognitive impairment, here defined as an IQ of less than 70 (Odding et al., 2006). It is more likely in children with severe physical disability (Novak et al., 2012), affecting almost 98% of these children

(Odding et al., 2006). Learning disability affects 60% of children with unilateral impairment, and its likelihood increases when epilepsy is a co-morbidity (Odding et al., 2006). Severe intellectual disability (of IQ less than 50) is entirely confined to children with all four limbs involved (Odding et al., 2006).

1.5.4.2.2 Epilepsy

There is a prevalence of 22 – 40% of epilepsy in children with cerebral palsy (Odding et al., 2006). Epilepsy is least common in children with only lower limb involvement (14%), and with dyskinetic or ataxic cerebral palsy (Odding et al., 2006). There is a strong association with cognitive impairment (Odding et al., 2006), and with severity of physical disability (Novak et al., 2012). Novak et al (2012) find that it is most common in children with all four limb involvement.

1.5.4.2.3 Visual and hearing problems

Visual impairment is most common in children with severe motor disability that is caused by spasticity (Novak et al., 2012), and the most severe visual disability or blindness occurs in 10% of children, half of whom have all four limb involvement (Odding et al., 2006). However, up to 71% of children may be affected by a visual impairment (Odding et al., 2006). Hearing impairments are less common (about 4% of children), but are most likely in those children with more severe physical disability (Novak et al., 2012).

1.5.4.2.4 Speech.

Speech impairment (usually dysarthria) can affect any children with cerebral palsy, but it is by far the most common in dyskinetic cerebral palsy, with 95% of children affected (Odding et al., 2006). However, up to 85% of children with spastic cerebral palsy with all four limbs involved have speech impairment (Odding et al., 2006), and a third of all children with cerebral palsy have some speech impairment (Novak et al., 2012).

1.5.4.2.5 Pain and musculoskeletal problems.

Chronic pain is reported in over a quarter of adults with cerebral palsy, compared with 15% in the general population (Odding et al., 2006), but it can affect children with cerebral palsy to a much greater extent (Novak et al., 2012). It does not relate to severity of disability, but is most likely to affect children with contractures (Novak et al., 2012), which are one of the secondary musculoskeletal problems associated with cerebral palsy, and is caused by spasticity (Thompson et al., 2005) and lack of opportunity for mobilising (Aisen

et al., 2011). Another common musculoskeletal problem experienced by many children with cerebral palsy is hip displacement and spinal deformity (scoliosis) (Novak et al., 2012). Both of these are increasingly likely with severity of disability, and they are most likely in those children with all four limbs affected and who are non-ambulatory (Novak et al., 2012). Scoliosis has an overall prevalence of 25% (up to 75% in those children with all four limbs affected (Koman et al., 2004)).

1.5.5 Prognosis

1.5.5.1 Mortality

Cerebral palsy is not, in itself, usually a fatal condition but there is evidence that it can lead to early death. One Californian study found the rate of mortality eight times greater in the cerebral palsy population than the general population (Strauss et al., 1999). In an earlier study, Strauss (1998) found that poor mobility and feeding difficulties were a strong predictor of early mortality. Severe cognitive impairment (O'Shea, 2008, Hutton and Pharoah, 2006), manual ability impairment and poor mobility (using an attendant-propelled wheelchair) is also a strong factor for predicting early mortality (Hutton and Pharoah, 2006).

1.5.5.2 Participation restriction

Beckung and Hagberg (2002) illustrate that severity of both mobility and manual ability are associated with a restriction of participation in school and social activities. Work prospects depend on reasonable cognitive ability, understandable speech, independent hand function and good mobility (O'Shea, 2008). In the UK, half of all disabled people are unemployed, and of those who do work, the majority are in very poorly-paid jobs (SCOPE, 2013). As more centres adopt a treatment philosophy based on the ICF the outlook for children and adults with cerebral palsy are expected to improve. Adaptations of the environment will increase participation in life situations e.g. lifts and ramps for access, speech assistive technology and other technological advances to overcome limited hand function e.g. eye gaze software, alternative keypads and other computer interfaces.

1.5.6 Summary of epidemiology, prognosis and demographics

Cerebral palsy is the most common cause of disability in childhood. Although essentially an unchanging neurological disorder, the impact of the cerebral disturbance impacts on the developing child's motor performance in a number of ways over their formative years. Both manual ability and gross motor ability are affected, and can be completely disabling

or barely noticeable. Commonly, there are associated neurological impairments e.g. epilepsy, cognitive or sensory impairments. Later in life, secondary musculoskeletal conditions often develop, worsening any existing disability. All of these cause activity limitation and participation restriction, impacting for example on the child's potential to participate in education and other important life situations, and limiting the potential for employment.

The potential for employment and involvement in life situations is, however, improving. Inclusion in UK schools and the workplace, for example, is driven by government legislation (Department of Education, 2003, Department of Education, 2004). However, there is still more that can be done. Although Osler identified the potential for functional gains from rehabilitation in children with cerebral palsy, children (and adults) were hidden away from public gaze, institutionalised in Victorian mental hospitals within our life time. In 1981, Joey Deacon (1920 - 1981) published his autobiography, *Tongue Tied*. *Tongue Tied* was an account of Joey's life throughout the half-century he spent in St Lawrence's Hospital (Ellis, 1982). Joey, who had cerebral palsy, was referred to the hospital aged only eight years old using language which, although in its day (1928) was not offensive, serves to illustrate the attitude visited on people with disabilities:

> *'a chronic and harmless lunatic, idiot or imbecile such as might lawfully be detained in an institution'* . (Ellis, 1982; page 485)

Joey's story served to help identify that greater efforts were required to recognise and implement services and adaptations to support and include people with disabilities. Today, children and adults with cerebral palsy are involved in schools, universities, the workplace and international sporting tournaments e.g. the Paralympics. Treatments and attitudes are evolving, and help for disabled people - from society, health care professionals and governments – is focussed more on treatments and environmental adaptations to enable full participation in life situations than at any time in history.

Addressing the cause of cerebral palsy is the subject of current research and is likely to involve prevention of prenatal infection, preventing premature birth and maintaining '*in utero'* oxygenation during and after an early birth. For now, the treatment of cerebral palsy focuses on addressing the impairments which are associated with the disorder in order to maximise the children's independence and participation and to enable them to lead as

normal a life as possible. The next section outlines these treatments with a focus on upper limb rehabilitation of children with spastic cerebral palsy.

1.6 Rehabilitation therapies for the upper limb of children with spastic cerebral palsy

1.6.1 Background and introduction

Spasticity affects up to 91% of children with cerebral palsy (Odding et al., 2006). Of these children, unilateral impairment is present in 21 – 40% and bilateral involvement (all four limbs) is present in 20 – 43% (Odding et al., 2006), therefore cerebral palsy causes the manual ability limitation in up to 83% of children with cerebral palsy.

Sir William Osler is the earliest medical practitioner in the literature to note the responsiveness to medical care of children with cerebral palsy (Longo and Ashwal, 1993) but orthopaedic surgeon Winthrop M. Phelps is credited with pioneering rehabilitation of children with cerebral palsy (Levitt, 2010). Phelps used a detailed classification system to describe cerebral palsy that formed the basis for 15 different treatments which were carried out by teams that included physiotherapists, occupational therapists and Speech and Language Therapy (Levitt, 2010). Such teams are still regarded as a fundamental component of present day NHS Trust Child Development Centres for children with developmental and motor delays (Mayston, 2004). Techniques developed by Phelps include repetitive passive, assisted and resisted movements of the upper limb that remain a central component of rehabilitation therapies today (Levitt, 2010), but until recently there was only poor evidence to support the use of these traditional rehabilitation therapies (Anttila et al., 2008). More recently, better quality trials have shown promising results for intense practice and repetition of functional activities (Gordon et al., 2006), and it has been suggested that one reason for the lack of evidence for the traditional rehabilitation therapies is that they performed too few repetitions and did not achieve an intensity of practice that is essential for functional change (Andersen et al., 2013). Therefore, current research into upper limb rehabilitation of children with cerebral palsy is directed towards investigating ways of encouraging children to practice repeated and intensive movements and functional activities of the upper limb. The basis and success of this research has influenced the development of the hypothesis of this thesis.

1.6.2 Aims of rehabilitation therapy

Motor impairment is the common denominator of all children with cerebral palsy. This impairment affects the child's ability to explore and interact with the environment that can result in potentially profound social, cognitive, emotional and developmental delays, perhaps with a lasting impact (Damiano, 2006, Levitt, 2010). The aim of rehabilitation therapy for children with cerebral palsy is to maximise each child's independent function (Levitt, 2010) and reduce activity limitation and improve participation in life situations; to prepare them for as normal a teenage and adult life as possible (Bobath and Bobath, 1984). The emphasis is on improving the quality of life (Mayston, 2004).

1.6.3 Neuroplasticity

It is said that practice makes perfect. On one hand, this applies to intellectual tasks such as navigation in the wilderness, games requiring tactical nous such as chess or draughts and indeed sports activities such as tennis; on the other hand, it applies to social situations and other circumstances where experience indicates that a previous course of action may or may not be a good idea. These forms of learning place some reliance on memories of emotional discomfort and pain, perhaps caused by social embarrassment, an unintentionally extended and ill-equipped extension to a wilderness adventure or a humiliating defeat.

On the other hand, sustained repetition of a motor task results in improved performance of the task, each time the task is performed. To achieve a superior standard of task performance requires practice sustained over many hours and days or weeks. There is the possibly apocryphal story about a famous golfer who scored, upon request, three consecutive holes-in-one on a training golf course. When reporters suggested that the golfer was lucky, the golfer remarked that he noticed how the harder he practiced, the luckier he became.

The successful golfer's swing, the tennis champion's serve and tennis stroke, the free-kick superiority of David Beckham and the goal-kicking records of Jonny Wilkinson owe much to hours of practice. The same is true of the virtuosity of a musician such as violinist Itzhak Perlman. There is evidence which suggests that sustained repetitive practice, causing repetitive sensory feedback and motor stimuli, drives an observable physiological response

and neuroanatomical changes. It is this physiological response and these neuroanatomical changes which underpin the current trend in rehabilitation therapy for children with cerebral palsy (Aisen et al., 2011, Nudo, 2003). The physiological response and the neuroanatomical changes stimulate the capability of neural tissue in the cerebral cortex to adapt, re-organise and form or strengthen synaptic connections in response to the external sensory and motor behavioural stimuli (Aisen et al., 2011). It provides the scientific rationale for the potential of repetitive functional movement to overcome impairment of movement (Nudo, 2003), suggesting that the areas of the motor cortex that map to body parts (see Figure 1-5) are adaptable throughout life, changing according to the amount they are stimulated (Scrutton, 2004).

Figure 1-5. The motor homunculus: a representation of the areas of the human body.

Figure 1-5a. The motor homunculus mapped to motor cortex.

Figure 1-5b. How a man would look if his body parts grew in relation to the areas of motor cortex areas that control them.

This phenomenon is strikingly illustrated in the MRI scans of string players whose left hand fingers are represented by a larger cortical area than that of non-string players, and that the area pertaining to the thumb (the least used digit in the fretting hand of string players) is the smallest area represented (Elbert et al., 1995). The hypothesis behind new experimental therapies therefore is that increasing the number of repetitive functional movements promotes synaptic generation and new neural pathways, establishing movement patterns which are essential for functional achievements e.g. walking, reaching, or fretting patterns on the violin. The question is whether or not this repetitive and intensive training of children with neurological damage results in reduced activity limitation and wider participation in life situations.

1.6.4 Development of rehabilitation therapies

Brain lesions can result in impaired movement and cause other impairments such as spasticity that further restrict the repetition of functional movement (Galea, 2004). This limits stimulation of synapses and re-organisation of the motor cortex that establishes neural pathways for controlling movement and increasing functional capability (Levitt, 2010). Methods to encourage repetitive or intensive movements are the basis of a range of new experimental rehabilitation therapies for which evidence is promising (Boyd et al., 2013, Green et al., 2013, Hoare et al., 2007a, Sakzewski, 2012). As long ago as the middle of the 20th century Phelps initiated treatment therapies using assisted and active movements supported by therapists (Levitt, 2010). Subsequently, treatments incorporating many of these movement therapies were developed e.g. Proprioceptive Neuromuscular Facilitation (Levitt, 2010) and the Bobath concept (Neurodevelopmental Treatment (NDT); Bobath and Bobath, 1984). NDT/Bobath is the most commonly-used concept in the UK (Mayston, 2004); it focuses on what is termed normal development, avoiding abnormal movement and postures, and establishing quality of movement (Levitt, 2010, Mayston, 2004, Miller, 2007) based on functional activities (Miller, 2007). But Scrutton (2004) suggests that normal movement is not the aim of rehabilitation therapy. He points out that humans adapt their movements to account for changes in environment e.g. astronauts on the moon "space-hop" rather than walk as on earth (Scrutton, 2004), and a person's gait is substantially different when wading through a swimming pool or across a fast-flowing stream. These movements are not 'normal' but are an adaptation to achieve a more effective functional movement. The emphasis on 'normal' has no foundation in research (Damiano, 2006), and therefore Scrutton (2004) suggests that the aim for children with cerebral palsy is to achieve *functional* movement over quality or "normality" of movement. In support of this Gordon et al. (2006) note that in contrast to NDT, rehabilitation therapies based on principles of intense repetition and practice of functional movements have good results. In fact, the quality of evidence in support of NDT is of such low quality that Novak et al. (2013) call for it to be discontinued as a functional rehabilitation therapy, although Andersen et al. (2013) suggest that the lack of evidence may be due to the lack of intense repetitions and practice. Further support for these ideas is offered by Boyd et al. (2001) who propose that a reason for greater success in lower limb rehabilitation is that the affected limbs are functional in a consistently repetitive manner, but that children with unilateral upper limb impairment usually disregard their affected upper limb, finding a way to perform even bilateral activity with their non-affected arm and performing unilateral tasks entirely with their unaffected side.

The issue of unilateral and bilateral upper limb use has influenced the development of measurement of upper limb activity limitation and the focus of research into rehabilitation therapies. There are a number of issues to consider. Firstly, with a few rare exceptions of truly ambidextrous people, each person has a natural preference for using their left or right hand limbs. The preferred limb is more commonly used for unilateral activities e.g. writing, cleaning teeth, and gains a better level of motor proficiency at these activities.

Naturally, children with cerebral palsy prefer to undertake unilateral activities with their less affected or non-impaired upper limb. This raises the question of why therapists focus on improving unilateral functional use of the impaired arm. For example, one would never consider targeting Roger Federer's non-dominant hand for improving his ability at tennis. Using another example, as one of the most gifted and athletic professional footballers in British history Ryan Giggs has won more trophies than any other footballer and has played at the highest level for over 20 years, yet he is criticised by some for being over-reliant on his dominant left foot. Given his ability and success, it is arguable whether Giggs should have spent longer trying to develop greater ability on his non-dominant right foot. These examples illustrate that unilateral activities are naturally and realistically performed by the preferred limb, usually the non-affected hand in unilaterally-impaired children, and that focusing on the impaired arm is only necessary in activity limited by impaired bilateral hand use.

In the case of unilateral upper limb impairment, functional limitation can be assessed by the Assisting Hand Assessment, which evaluates the effectiveness of the affected hand in bilateral play activity (Hoare et al., 2013). But some measures are designed to measure unilateral arm and hand use e.g. the Melbourne Assessment of Unilateral Upper Limb Function, and other activity measures such as the ABILHAND-kids questionnaire includes unilateral activities. This again raises the question, why would therapists want to measure the upper limb activity limitation of the affected arm in unilateral activities? To illustrate the point, some parents completing the ABILHAND-kids, when asked to describe how easy their child finds each unilateral activity, are at a loss to answer because it depends on which arm the child uses. This is not a question which would cause the same confusion

when asked of a non-disabled child, or if one were to ask Roger Federer if he could play tennis.

The most recent research into rehabilitation therapies of the upper limb include investigations into the benefits of intense training of both unilateral and bilateral training on activity limitation, and into the effects on this training on neuroplastic changes (Andersen et al., 2013).

1.6.4.1 Research into rehabilitation therapies

There is a poor evidence base for traditional methods of rehabilitation therapy, with physiotherapy suffering from a bad reputation for research until recent years (Anttila et al., 2008). There are a number of reasons for this.

First of all, a number of systematic reviews have found poor methodological quality in efficacy trials (Anttila et al., 2008, Novak et al., 2013, Sakzewski et al., 2009, Boyd et al., 2001). Problems include poor reporting, small sample sizes, lack of control groups, underpowered groups or no power calculations, and a lack of blinding. Sakzewski et al. (2009) suggest that poor adherence to NDT programmes (partial completion of the therapy on 18% of days) may also negatively influence the evidence for the benefits that might be gained from more diligent participation in rehabilitation therapies.

A second problem, and one which remains a problem for current research, is the heterogeneity of the cerebral palsy population (Anttila et al., 2008, Boyd et al., 2001), especially in the movement disorders involving the upper limb in children with cerebral palsy (Boyd et al., 2001). With a heterogeneous group of participants, it is difficult to generalise the results to narrower classifications of the population, and the large variation in the presentation of children makes meta-analyses of trials impossible. There is also the potential that certain therapies might have a greater impact on a particular, narrowly-defined sub-group with cerebral palsy, so that the results of studies on a more heterogeneous sample might not identify the changes in activity limitation in that sub-group. For example, Damiano (2006) suggests that intervention as early as infancy is essential because there is the potential of increased plasticity in the infant's brain, as

evidenced by the observed natural destruction of unused neurons and the strong reinforcement of developing neural pathways in use at that age (Damiano, 2006). Leaving rehabilitation beyond infancy permits reinforcement of those neural pathways that control and drive the abnormal and non-functional movements and allows the child's abnormal and less functional movements to develop.

A third problem is the use of inappropriate measures, or use of measures with no evidence of validation in children with cerebral palsy (Hoare et al., 2007a, Sakzewski et al., 2009, Eliasson et al., 2005). The use of such measures to evaluate changes in activity limitation affects the validity of the results and limits the impact of research findings. Systematic reviews of upper limb activity measures suggest that only three possess adequate psychometric qualities for use with children with cerebral palsy (Gilmore et al., 2010, Greaves et al., 2010, Harvey et al., 2008, Klingels et al., 2010):

- The Melbourne Assessment of Unilateral Upper Limb Function (MAUULF) ;
- Assisting Hand Assessment (AHA);
- ABILHAND-kids.

The importance of appropriate outcome measures to evaluate the outcomes of this research study prompted a systematic search and critical appraisal of available measures in order to identify the measure which was valid, reliable and responsive for use with children with cerebral palsy. The conduct, results and conclusions of this appraisal are described below in Chapter 3 "Measuring activity limitation and kinematics in children with cerebral palsy"; however, the appraisal identified problems with the interpretation and use of ordinal outcome scores (see paragraph 3.2.2 "The problem with ordinal outcome scores" on page 89) and noted problems even with the MAUULF, the AHA and the ABILHAND-kids. For example, the MAUULF includes among its items a large proportion that relate to Body Functions (Hoare et al., 2011) and produces ordinal level outcome scores, so its usefulness as an activity outcome measure in research is questionable. The AHA (Krumlinde-Sundholm et al., 2003) and the ABILHAND-kids (Arnould et al., 2004) produce outcome scores that can be transformed into interval level scores. Both are now in frequent use as primary outcome measures in modern research. The AHA is a test of bimanual performance, evaluating the use of the affected arm and hand in bimanual play activities; the ABILHAND-kids is a questionnaire of 21 mostly bimanual activities of self-care (Aarts et al., 2010). However, the AHA is costly in terms of purchase, training and scoring (Gilmore et al., 2010, Gordon et al., 2006, Greaves et al., 2010, Krumlinde-Sundholm et al., 2007),

while the ABILHAND-kids psychometric testing was performed on a sample of children with cerebral palsy very few of whom experienced very limited activity of the upper limb (Arnould et al., 2004).

Other measures that have been used with children with cerebral palsy include the Canadian Outcome Performance Measure (COPM; Cusick et al., 2007, Verkerk et al., 2006) and Goal Attainment Scaling (GAS; Steenbeek et al., 2007) but these measures calculate outcome scores using ill-founded mathematical procedures on ordinal data (Grimby et al., 2012, Stucki et al., 1996, Tennant, 2007), and because they are not standardised they are unsuitable for group comparisons (Tennant, 2007). A limited understanding of the nature of ordinal data can confound the integrity of the results of even well-designed trials. For example, Wallen et al (2007) used the GAS and COPM as the primary outcome measures but despite these producing ordinal outcome scores, the authors calculated the scores as means and standard deviations, exacerbating the mathematical inconsistencies of their outcome score generation. The COPM and GAS scores from this study showed a statistically and clinically significant improvement in the participants' perception of functional performance, but the study's other measures (the MAUULF, the Quality of Upper Extremity Skills Test (QUEST) and Paediatric Evaluation of Disability Inventory (PEDI)), all commonly used in both research and clinical practice with children with cerebral palsy, showed no changes in outcome scores, suggesting that only the COPM and the GAS are responsive to changes that demonstrate efficacy of experimental treatments or treatment. The COPM shows responsiveness in a number of other clinical studies compared to measures that present a standardised list of activities (Sakzewski et al., 2007). However, the COPM and the GAS items are activities (up to five) that are selected at baseline following a detailed assessment of activity limitations, and each item is given a score to indicate the child's capability; then, once the activity limitation has been addressed by therapy or rehabilitation, it is re-scored. Items are therefore selected with a likelihood of some improvement in scores, because they are selected as potential areas where therapy will have an impact. Additionally, items might include, for example, 'to increase wrist extension', or 'to reduce spasticity of the elbow flexors'; in this case, a change of these impairments which cause a change in outcome scores that imply improved activity limitation.

Recent research has addressed to some degree methodological flaws, with encouraging results for newer experimental rehabilitation therapies such as Constraint Induced Movement Therapy and bimanual training (Aarts et al., 2010, Hoare et al., 2013, Novak et al., 2013). These have focussed on ways to elicit intensive practice of functional use of the impaired arm.

1.6.4.2 Constraint Induced Movement Therapy and Bilateral Training

Andersen et al. (2013) suggest that insufficient practice and too few repetitions of functional movements may explain the lack of evidence for traditional movement rehabilitation therapy and in support of this, Boyd et al. (2013) finds increasing and consistent evidence that intensive repetition of functional movement has the potential to overcome the developmental disregard and improve activity limitation of the impaired limb of unilaterally-impaired children. Constraint Induced Movement Therapy (CIMT; Taub, 2004, Taub et al., 1998, Taub and Uswatte, 2003, Taub et al., 1999) places a restrictive element (e.g. a glove, splint or arm sling) on the non-affected arm of children with unilateral impairment to compel the child to perform activity primarily with their impaired arm with the aim of reducing activity limitation (Sakzewski et al., 2009). An important feature of CIMT is high intensity functional activity, forcing practice of the activity using the affected arm (Hoare et al., 2007a). CIMT programmes can result in cortical reorganisation on MRI imaging (Sterling et al., 2013). Hoare et al. (2007a) conclude that CIMT should be regarded as an experimental therapy until better quality trials are conducted, with the use of valid and reliable outcome measures essential, but a more recent review which includes trials performed subsequently strongly recommends CIMT to improve functional use of the impaired hand of unilaterally-impaired children with cerebral palsy, although quality of evidence remains moderate (Novak et al., 2013).

Arguments against CIMT include its intensive and invasive nature and lack of practice on bilateral activity (Boyd et al., 2013). Intensity of CIMT programmes may account for a lack of adherence (limited to 57%; Sakzewski et al., 2009) and drop-outs from treatment (Gordon et al., 2006, Wallen et al., 2011) which may in turn negatively affect the outcomes of research trials or clinical practice. Also, CIMT concentrates on unilateral activity but independence in daily functional activities requires two-handed skills (Sakzewski et al., 2011b, Gordon et al., 2007), which may be reflected by the number of bilateral goals formulated by parents in individualised outcome measures such as the COPM and the GAS (Sakzewski, 2012, Wallen et al., 2011). This supports arguments for a therapy which

includes bilateral activity (Charles and Gordon, 2006). Charles and Gordon (2006) further suggest that the key to functional gain is bilateral practice rather than constraint and that the aim of rehabilitation is not increasing unilateral arm use (of the affected arm) but improving coordination of both hands, thus improving activity. Aarts et al. (2010) disagree with Charles and Gordon (2006), arguing that children with unilateral impairment possess an element of developmental disregard that interferes with bilateral coordination. Furthermore, evidence from two trials suggests that CIMT shows a positive and clinically significant effect on bilateral hand use (Eliasson et al., 2005, Hoare et al., 2013), and that this clinically significant change is comparable with bilateral therapy when both are used with botulinum toxin (Hoare et al., 2013). Both of these trials used the AHA, which evaluates bilateral use and has strong psychometric properties, for their primary outcomes. However, intensive bilateral training may be particularly useful when the activities to be practiced involve a level of grasping too skilful and coordinated for the affected hand or when the child does not tolerate the restraint (Andersen et al., 2013) .

Charles and Gordon (2006) emphasise the inclusion within intensive rehabilitation therapy such as bilateral training or CIMT of specific pre-defined goals. This is supported by a recent systematic review of interventions for children with cerebral palsy by Novak et al. (2013) which finds strong evidence for goal-directed training. However, one such high quality study includes their own trial in which they suggest that home programmes of occupational therapy show significant improvements in functional goals using the COPM (Novak et al., 2009). The authors report a mean clinically significant change of 2.4. In fact, although a clinically significant change in COPM scores is a minimum of 2 (Law et al., 2005), the results state a confidence interval of 0.7 to 4.2, which suggests that there may be no clinically significant change at all. There are a number of possible reasons why the outcome did not show a significant change (heterogeneity of children, incorrect use of statistics (using parametric analyses on non-parametric data), heterogeneity of other interventions within the groups) but the main difference between this study and others investigating CIMT and bilateral intensive training is the intensity of intervention. In their study, Novak et al. (2009) found that parents conducted their home therapy programme for a mean of 16.5 minutes on 17.5 days per month (mean 288 minutes per month), a fraction of the time spent in intensive therapy e.g. mean 7,200 minutes per month (Charles and Gordon, 2006) or 2,160 minutes per month (Aarts et al., 2011). This suggests that intensity of repetition may be the most defining factor associated with improvement (Novak et al.,

2009). Another interesting study which suggests that intensity of bilateral rehabilitation rather than functional goal-focused rehabilitation is the crucial component used a magician training camp, in which children with cerebral palsy learned to perform a magic trick that required bilateral hand use (Green et al., 2013). This study involved a two-week (6 hours per day) training camp which focussed entirely on practicing and learning magic tricks, on the design and building of props for the tricks and included time for bilateral play activities (Green et al., 2013). It is also notable for its range of ages and disabilities in the children recruited. Limitations include that the small number of participants were a convenience sample and there was no control group. However, it was correctly powered, assessors were blinded and the study resulted in significant improvements in AHA scores. This study supports the suggestion that intensity of bilateral practice on fun activity is more important than rehabilitation with an emphasis on specific functional goals. It is worth noting that no children withdrew from either this programme or from the programme with pirate-themed activities (Aarts et al., 2010), whereas withdrawals appear common in "standard" CIMT studies (Choudhary et al., 2013, Eliasson et al., 2005, Gordon et al., 2006), with commitment and intolerance of the intervention cited as the reason for withdrawal of participants (Gordon et al., 2006).

The optimal dose of these therapies has not been established (Andersen et al., 2013). In a case study of CIMT, an eight-year-old boy showed cortical reorganisation and functional benefits (measured by a measure of very low psychometric properties) after 3 weeks of continuous casting with only one hour per week of occupational therapy (Sutcliffe et al., 2007), but the high quality evidence for CIMT comes from studies with up to 30 hours per week of intensive training (Aarts et al., 2010, Hoare et al., 2013, Sakzewski et al., 2011b). A recognition that these intensive programmes are a substantial commitment for families and for rehabilitation teams has led to trials of a modified CIMT programme which has much less intensive training. These trials have showed promising results (Aarts et al., 2010, Eliasson et al., 2005, Wallen et al., 2011). Modified CIMT (mCIMT) involves only a 2-hour period of restraint for the affected limb for seven days a week for two months (Eliasson et al., 2005), and can incorporate a daily home programme of 30 minutes per day to be conducted by parents (Wallen et al., 2011). One advantage of this is that the constraint can be worn in any environment, without disruption to education and social activities e.g. home or school rather than in a group therapy programme or therapy clinic, a factor that is

increasingly recognised as important for beneficial rehabilitation outcomes (Eliasson, 2005, Scrutton, 2004).

Andersen et al. (2013) note that while both unilateral and bilateral training approaches show benefits in bilateral ability, CIMT is the best approach for improving unilateral impairment and activity limitation while intensive bilateral training is best used for improving bilateral coordination and achieving functional goals. It is recognised that functional activities tend to be bilateral and that parents emphasise that their children show greatest limitation with such activities, but when a child shows marked disregard for their affected limb then focussing on the affected limb may be advantageous as a precursor to bilateral rehabilitation (Aarts et al., 2010). In a powered RCT to evaluate the benefits of CIMT followed by bilateral training, Aarts et al. (2010) found a substantially greater improvement in the experimental group over the control group. Each group took part in eight weeks of rehabilitation therapy at the rehabilitation centre, with the control group having twice-weekly therapy sessions (one and a half hours per week) and the experimental group taking part in pirate-themed sessions (nine hours per week). The experimental group received six weeks of CIMT followed by two weeks of bilateral training that emphasised goal-directed play and self-care activity. Both groups' parents were encouraged to promote affected arm use at home (additional recorded stimulation of the affected upper limb: control group = eleven hours, experimental group = three hours). Both groups showed an improvement in AHA and ABILHAND-kids scores but the experimental group improvement was two and a half times greater (AHA score) and seven times greater (ABILHAND-kids score). This was a well-conducted study but potentially the experimental group received far more active rehabilitation and the pirate theme may have been more stimulating for this group. Given that previous studies have shown the potential for the benefits of both CIMT and intensive bilateral training compared to home programmes and conventional therapy, the results of this trial might have been more meaningful if they had used a CIMT or an intensive bilateral training intervention as a comparison group, but as yet there are no trials which investigate the respective benefits of CIMT-bilateral training intervention, CIMT or bilateral training intervention over each other. However, there are suggestions that the timing and nature of each intervention has serious implications for rehabilitation outcomes. These suggestions are based on evidence in adult and animal studies that motor function is adversely affected through the influence of ipsilateral corticospinal connections, shown in Figure 1-3 on page 20. At birth, as well as

the contralateral neural connections, there are a large number of ipsilateral projections that undergo major neuroplastic changes in that they withdraw, and contralateral projections are reinforced (Kirton, 2013). Kirton (2013) suggests that the ipsilateral projections are the reason that term babies with unilateral lesions do not present with hemiparesis, and recognises that their existence might imply strong potential for good function and activity. But Kirton (2013) proposes instead that this is "maladaptive" (page 85), a final "better than nothing" option for the body's natural adaptive healing and developmental capability. Additionally, neurological connections *between hemispheres* may negatively influence plasticity, with the non-lesioned hemisphere inhibiting the capability of the lesioned hemisphere to develop or reinforce motor projections to the contralateral hand, and thus increasing the counterproductive ipsilateral connections which are associated with poorer hand activity (Kirton, 2013). Andersen et al. (2013) recognise that these projections may be influenced by rehabilitation, suggesting that CIMT may reduce the inhibition of the unlesioned hemisphere over the lesioned hemisphere, and promote plastic development and reinforcement of contralateral corticospinal projections. This promotes the idea of a two-stage rehabilitation programme, beginning with unilateral training to overcome the effects of neuroplastic maladaptation, developmental disregard and poor manual dexterity and followed up with bilateral training to promote functional goal-directed activity and bilateral coordination (Aarts et al., 2010, Andersen et al., 2013).

1.6.5 Botulinum toxin treatment for spasticity

Spasticity affects over 90% of children with cerebral palsy (Odding et al., 2006). Spasticity affects numerous functional activities including washing, dressing, and picking up items (Graham et al., 2000). Untreated, spasticity can also lead to muscle shortening and soft tissue contractures which have an additional detrimental impact on functional activity (Graham et al., 2000). Treatment of spasticity has therefore always played an important part in rehabilitation and the medical management of children with cerebral palsy. The use of botulinum toxin is a recent addition to pharmacological management of spasticity via localised injection (Graham et al., 2000). It works by blocking the release of the neurotransmitter acetylcholine at the motor endplate, preventing muscle activity (Barnes, 2003). It is therefore a focal treatment, targeting specific carefully-selected muscles. There is high level evidence that botulinum toxin produces effective upper limb functional benefits only when used alongside rehabilitation therapy (Hoare et al., 2010, Hoare et al., 2013).

1.6.6 Summary of rehabilitation therapies

There is moderate to good evidence in support of the hypothesis that repetitive and intensive practice of unilateral and bilateral activity improves functional independence and reduces activity limitation in children with cerebral palsy. A number of factors that are likely to influence the functional benefits of CIMT and bilateral training therapies, such as age and level of disability, remain to be explored but the intensity of practice and number of repetitions are critical elements. Investigations into neural pathways and neuroplasticity suggest that the best use of intense practice and repetition is to focus initially and at an early age on unilateral training (CIMT) followed by goal-directed and other stimulating bilateral training of high intensity and repetition. There are also promising results in trials combining CIMT and bilateral training with spasticity treatments, such as botulinum toxin which is known to have a beneficial effect on daily activity when combined with rehabilitation therapy. Some children or families appear to find the intense nature of the rehabilitation therapies challenging but adherence to the programmes can be improved by practicing unusual and stimulating skills or using imaginative and innovative themes. For these reasons, there has been increasing interest in the use of virtual reality, video-gaming and robotics to supplement rehabilitation programmes or enhance the benefits of adjunctive treatments such as botulinum toxin treatment.

1.7 Summary of Chapter 1

Cerebral palsy is a common cause of disability in childhood, causing upper limb activity limitation in up to 83% of children. This can affect the opportunities for exploration and social interaction that are essential for typical development of children, causing a detrimental impact on the child's participation in education and other important life situations, and limiting the potential for employment in adulthood.

Current treatment of cerebral palsy focuses on addressing the impairments which are associated with cerebral palsy in order to maximise the children's independence and participation and to enable them to lead as normal a life as possible.

There is moderate to good evidence that repetitive and intensive practice of unilateral and bilateral activity reduces activity limitation and increases independence. Some research suggests that high intensity unilateral training (CIMT) followed by bilateral training produces greater benefits. There are also promising results in trials combining CIMT and bilateral training with botulinum toxin treatment. To promote interest in these therapies,

the intensive nature of which can be challenging, the use of virtual reality, video-gaming and robotics to supplement rehabilitation programmes has been proposed.

2 The use of virtual reality, robotics and computer games to supplement rehabilitation programmes or enhance pharmaceutical interventions

2.1 Background and overview of gaming technology and virtual reality to support rehabilitation of children with cerebral palsy

There is growing evidence that repetitive and intensive practice of unilateral and bilateral activity promotes activity in children with cerebral palsy, and that concordance with rehabilitation programmes can be improved by practicing unusual and stimulating skills or using imaginative and innovative themes. The potential for modern technology to support these types of therapy has resulted in mounting interest in the use of robotics, gaming technology and virtual reality to motivate children to engage with rehabilitation programmes (Fasoli et al., 2010, Krebs et al., 2009, Meyer-Heim and van Hedel, 2013, Qiu et al., 2009). Virtual reality and gaming technology can establish child-friendly, pleasurable, captivating, motivating and interactive scenarios (Qiu et al., 2009, Sandlund et al., 2009). Virtual reality also provides safe environments for the simulation of real-life activities than could potentially promote development of skills that transfer to actual reality (Wang and Reid, 2011). Furthermore, Fasoli et al. (2010) and Sandlund et al. (2009) suggest that robotic-assisted therapy is more inclusive for children who have a severity of upper limb impairment that prevents their participation in CIMT/bilateral therapy sessions.

However, few studies to provide evidence for the usefulness of these technologies have been carried out (Meyer-Heim and van Hedel, 2013, Mitchell et al., 2012, Sandlund et al., 2009), and these consist of case studies, multiple case studies and uncontrolled studies. A systematic search to identify whether papers published more recently included better quality studies revealed a total of 413 papers of which only 18 evaluated the benefits of using virtual reality, gaming, assistive or robotic technology in children with cerebral palsy (see Table 2-1 below).

Table 2-1. Search strategy for articles that evaluate activity limitation benefits from the use of assistive technology

Databases searched: AMED, all EBM reviews, Embase, Ovid Medline
Searched on 20th January 2014

	SEARCH TERMS
1	Cerebral palsy
2	CP
3	Hemiplegia
4	Quadriplegia
5	Diplegia
6	Spastic*
7	unilateral
8	1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7
9	technology
10	Assistive technology
11	Rehabilitation technology
12	Computer games
13	Wii
14	9 OR 10 OR 11 OR 12 OR 13
15	Physio*
16	Occupational
17	Rehabilitation
18	Exercise
19	Strength*
20	Movement
21	Constraint
22	15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21
23	Function*
24	Activity
25	Participation
26	ADL
27	23 OR 24 OR 25 OR 26
28	8 AND 14 AND 22 AND 27
29	Remove duplicates

Gordon et al. (2012) carried out a feasibility study (no blinding and no controls) with the Nintendo Wii™ using seven children, one of whom dropped out due to parental disinterest. After six weeks of using the games system on two days a week (weekly total 90 minutes), the children showed a substantial improvement in gross motor movement, but

the authors note that a powered RCT would require 128 children altogether. Other studies also demonstrated feasibility of such approaches but again used only single cases or case series (Borggraefe et al., 2010, Deutsch et al., 2008, Fluet et al., 2010, Golomb et al., 2010, Green and Wilson, 2012, Qiu et al., 2009). Golomb et al. (2010) used a home-based device that was networked through both the rehabilitation centre and the research establishment, and used a commercial games station (PlayStation3™). Golomb et al. (2010) found an improvement in two out of the three participants, the third of which had a faulty device that limited the number of days on which they played the device; however, each played for a mean of over twenty minutes a day and the children played on between 36 and 67 days in total. Although the results did suggest a subjective change, the measures (the Bruininks-Oseretsky Test of Motor Proficiency (BOTMP) and the Jebsen–Taylor Test of Hand Function (JTT)) were not suitable for assessments of this type (see section 3.3 below and Table 3-5 on page 117 for review and characteristics of measures used to evaluate upper limb activity limitation of children with cerebral palsy). Green and Wilson (2012) report on four out of eight children in their study, finding that children engaged with the virtual reality system and showed potential for improvements; they also note that the optimal dose of using such technology is uncertain, and assert that individualised measures such as the Canadian Occupational Performance Measure (COPM) are useful for such studies in spite of their recognised problem with non-standardisation. Fasoli et al. (2008) suggest that using robotic technology to facilitate repetitive, goal-directed training in combination with botulinum toxin produces the same results as a program of occupational therapy combined with botulinum toxin treatment, but this is again a single case study and the outcome measures are not accurate or appropriate for children with cerebral palsy.

Winkels et al. (2013) carried out an exploratory study using the Nintendo Wii™ on a convenience sample of 15 children with mild to minimal upper limb impairment. There was no control group, and the children used the Wii™ for 30 minutes per session twice a week for six weeks. There was no change in activity limitation as measured by the MAUULF and the ABILHAND-kids. The children reported pain in their arm in the early stages of the study which may have been caused by unaccustomed exercise of the limb, suggesting that an hour a week may have caused some exercise effects but not enough to cause functional benefits. These results might also suggest that these particular measures lack responsiveness. The lack of responsiveness and validity of outcome measures is reported by a number of researchers (Palsbo and Hood-Szivek, 2012, Qiu et al., 2009, Sandlund et

al., 2009), supporting a conclusion that is common among investigators evaluating the benefits of botulinum toxin, CIMT and bilateral training on activity limitation e.g. Hoare et al. (2010).

The main drawback with the use of existing technology and games is that they are not designed or produced to target movements and capabilities of disabled children, while the development of customised technology and software is expensive and time-consuming (Wang and Reid, 2011). Wang and Reid (2011) further suggest that, for rehabilitation purposes, virtual reality (and, by association in this context, some gaming activities at least) should mirror real life to stimulate the type of activity required to achieve reduction in activity limitation. They propose that research into assistive and virtual reality technology should focus on three key areas (Wang and Reid, 2011):

1. the development of affordable computer-assisted rehabilitation or adapted commercial games systems;
2. the use of virtual reality as an intervention that could potentially benefit a range of impairments e.g. coordination and cognitive impairments;
3. the need to measure benefits of technology ("skill transfer", page 16).

In parallel with these developments in the use of technology for rehabilitation purposes, a team at the University of Leeds developed its own assistive computer gaming technology, and investigated its potential to improve activity limitation of children with cerebral palsy.

2.2 Design, construction and feasibility trial of prototype home-based assistive joystick and computer games (NIHR-funded study, grant ID G006)

2.2.1 Background – initiating the study

The primary outcome of this NIHR-funded project was to develop a prototype of a computer game system that would assist children with cerebral palsy who had upper limb movement impairment to undertake therapeutic movements at home. It was necessary to construct two separate components:

- A powered joystick, adjustable to suit an individual child's range of upper limb movement and ability to grasp.
- Appropriate computer games that encourage therapeutically useful repetitive reach-retrieve arm movements. The gaming software also had to control the assistance provided by the joystick and measure kinematic aspects of its movement.

Secondary outcomes were:

- the establishment of mechanisms for involving children, teachers and parents in the design of this and future therapeutic equipment;
- to establish the feasibility of using the system to undertake supplementary exercises at home; and
- to evaluate changes in activity limitation and arm kinematics following use of the games system.

Recognising that a strong multidisciplinary team is essential (Meyer-Heim and van Hedel, 2013), a team consisting of engineers, designers, paediatric therapists, rehabilitation consultants, academics and psychologists was assembled at the University of Leeds to conceptualise, design, build and test a motor-powered assistive computer game for children with cerebral palsy. The games would encourage repetitive reach-retrieve movements and aimed to engage children with cerebral palsy to undertake repeated game play. It was intended to evaluate the potential of the assistive games system for home use and evaluate change in upper limb activity limitation through a feasibility study in which children were asked to play the games system daily in their homes (Weightman et al., 2011).

2.2.2 Design of games and hardware components

It was recognised that an essential but oft-neglected component of design is inclusion of the end user, particularly when the end user is a child (Druin, 2002), so an 'end user' group was established consisting of children with cerebral palsy whom were approached through local paediatricians and occupational therapy teams. Included within the user group were a number of wheelchair users, which was essential to ensure that a system was developed which did not exclude this population of children with cerebral palsy.

An important consideration was the desire to avoid labelling of the technology as a medical-type device 'for the disabled'. Therefore, some of the design and feedback procedures were carried out at the children's schools and the children were asked to invite their classmates, none of whom had any disability, to participate in the school-based design activities. Our user-centred design and feedback sessions therefore included five children with cerebral palsy and 37 non-disabled classmates. It was also essential to establish how enjoyable the games were, how easy they were to set up independently and play, and whether the games system interface (the joystick) was user friendly. The main purpose of the user group was therefore to answer four questions (Weightman et al., 2009):

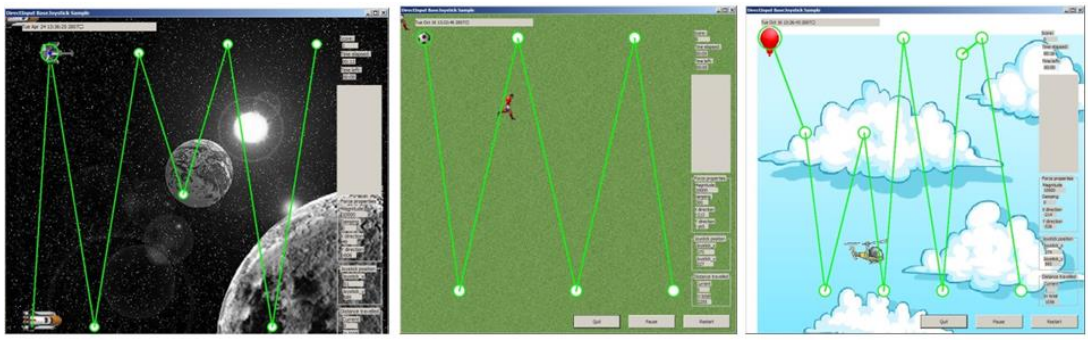
1. Did the children associate the games system with disability or view it as a medical device?
2. Was the joystick comfortable to use during and after a 15-min playing session?
3. With the games system set up, were the children able to independently set up and play games?
4. Did the children engage with the games?

For these user group meetings, a method and a 'tool kit' developed by previous researchers who included children as end user designers was adopted. The method was peer tutoring (Höysniemi et al., 2003), the primary aim of which was to address the third question but which also afforded an opportunity to observe the children and capture information relating to the second and fourth questions. Peer tutoring involved the instruction of a child on setting up and playing the games system, before an iterative process in which children instructed and showed subsequent children in turn how to set up and play the games system. Questionnaires and a 'Fun Toolkit' (e.g. use of a Smileyometer; Read et al., 2002) were developed which were used before the peer-tutoring sessions to address the first question, and then after the sessions to address the remaining three questions. Different versions of the joystick, joystick handles and their coverings (see Figure 2-1) were also evaluated by the children through use of questionnaires, with the emphasis on answering the first and second questions.

Figure 2-1. Potential handles for the gaming interface (joystick).

These procedures guided development of the games and two versions of the games system joystick, an adapted Microsoft Sidewinder gaming joystick. The games are shown in Figure 2-2.

Figure 2-2. The G006 games.



The games were all based on the same underlying principle and mode of operation, with only the graphics changing in an attempt to maintain children's engagement. The theme of each game was the movement of the character or sprite (e.g. a spaceship, footballer, or helicopter) to a series of appropriate targets (e.g. a docking station, a goal, or balloons), thus eliciting the required arm movements by appropriate arrangement of the targets (Weightman et al., 2011).

The joysticks are shown in Figure 2-3 below. The joystick on the left is to accommodate the situations where limited space is available for home use; the joystick on the right is

designed to provide greater torque (assistance) for children with a greater degree of spasticity or impaired movement.

Figure 2-3. The G006 joysticks. The joystick on the left is for situations when space is restricted. The joystick on the right can provide greater torque.

The amount of assistance provided by the joysticks can be easily adjusted through the gaming software.

2.2.3 The question of trunk restraint

The question of whether or not to provide specialist seating or trunk restraint was an important consideration for this study. There are suggestions that children taking part in task-oriented training who have trunk movement restraint imposed upon them show more improvements in upper limb kinematics than in children with no trunk restraints, though these improvements did not extend to functional improvements (measured by the Melbourne Assessment of Unilateral Upper Limb Function (MAUULF, see page 123)) (Schneiberg et al., 2010). However, most children within the user group were not provided with such supportive seating or trunk restraints by their clinicians, and such equipment might also restrict the use of the games system. One user group family who spent every weekend in a caravan suggested that they would not use such a restraint because it would impact negatively on the child's use of the games. It was decided that this study would not require trunk restraint other than that already supplied by the child's clinician.

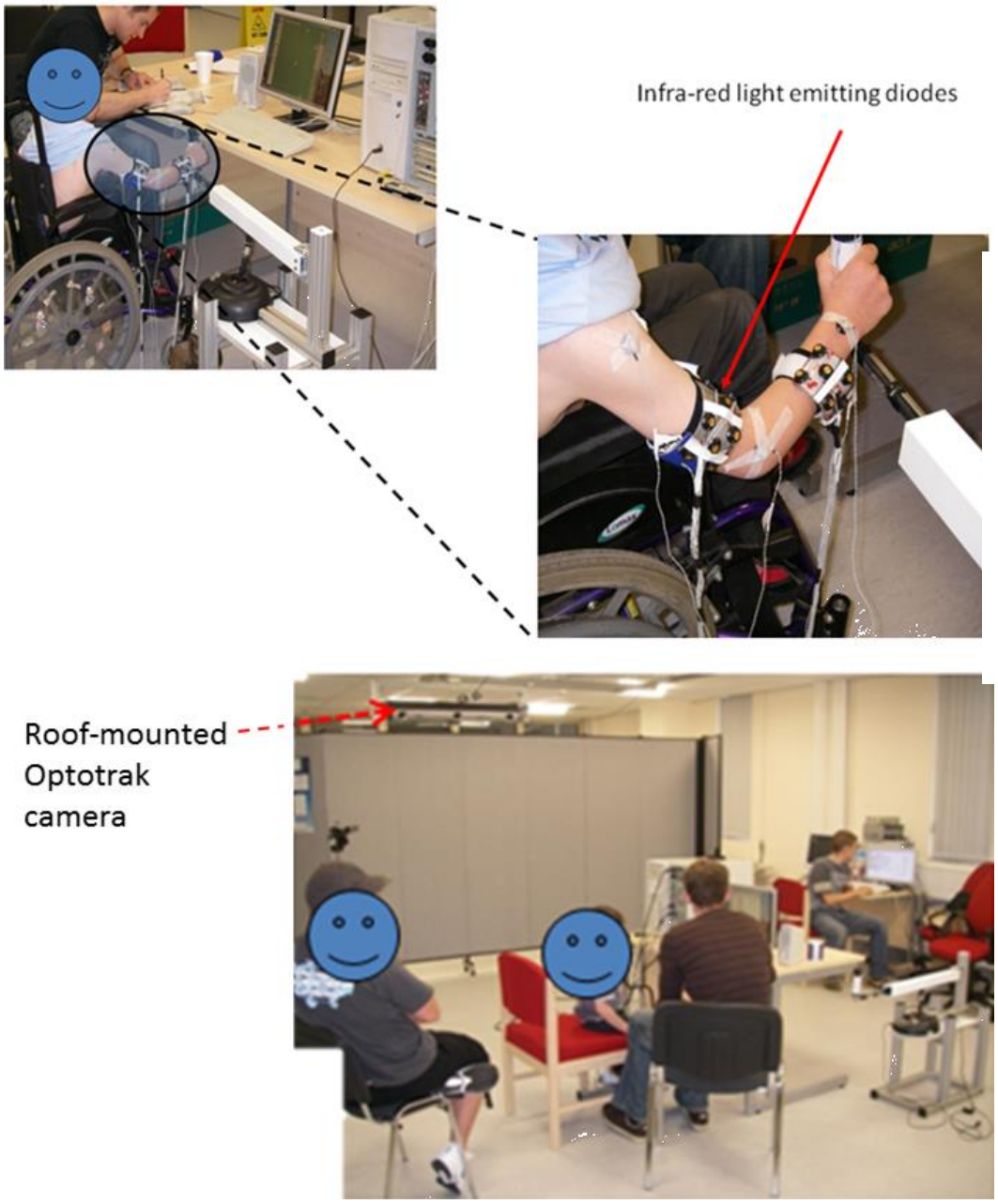
2.2.4 Feasibility study of home use (Weightman et al., 2011)

A total of 18 children with cerebral palsy aged 5–16 years old, (median 7½ years, 13 males) were recruited, six of whom were wheelchair users. The children were approached on our behalf by paediatricians and therapists from local paediatric teams. Some of the children were familiar with the games system and the research team because of their involvement with the user-centred design process described in the previous section (Weightman et al., 2009).

2.2.4.1 Outcome measures

All children had significant upper limb impairment of at least one upper limb (14 showed impaired right upper limb) as measured on the Movement Assessment Battery for Children (MABC; Henderson and Sugden, 1992) at which each child scored the lowest possible functional grade. The MABC is not validated specifically for children with cerebral palsy, and most of the MABC tasks were beyond the capability of the children, suggesting a floor effect for the MABC. The Canadian Occupational Performance Measure (COPM; Law et al., 2005) was used to evaluate changes in activity limitation for activities identified as limited for each child individually. Game use was an important metric to evaluate how well the games system engaged the children, and this outcome was evaluated with use of the same questionnaires and toolkit that was used in the design process (see previous section and Weightman et al., 2009). Finally, arm kinematics were evaluated at baseline and post-intervention through the use of an Optotrak Certus movement recording system, situated at the University of Leeds' Charterhouse Rehabilitation Technologies (CRT) Laboratory. Figure 2-4 below shows the movement recording system being used for the kinematic assessment of a child at the CRT Laboratory.

Figure 2-4. The Optotrak system in use at the CRT Laboratory to capture upper limb kinematics.



2.2.4.2 Feasibility study design

The games system was deployed to each of the children's homes for four weeks. A researcher visited the child's home at two weeks to review the settings of the game and adjust the assistance if necessary. All baseline measures were conducted at the CRT Laboratory, University of Leeds within a few days of deployment of the games system to the child's home.

2.2.4.3 Results

2.2.4.3.1 Usage

The children played the games for a varying amount of time, a median of 75 minutes per child over the four weeks (range 0.2–271 min). They practised a median of 606 outward movements (range 4–1684) and a median of 734 inward movements (range 5–2041). Two children used the games system for less than 30 min over the four-week period (Weightman et al., 2011). The results of the questionnaires showed that the children enjoyed playing the games but revealed a preference of the children for games which permitted friends and family to play together in dual user games.

2.2.4.3.2 Activity limitation

Changes in activity limitation measured on the COPM showed a statistically significant improvement. Median scores improved from 4.2 (range 1 to 5.6) to 6.0 (range 4.4 to 7.4, $p < 0.001$) (Weightman et al., 2011). The median change score for a clinically significant difference is 2.0 (Law et al., 2005), which was not achieved. Individual and median change scores are shown in Table 2-2.

Table 2-2. COPM scores for G006 (Weightman et al., 2011)

Participant	Initial COPM	Final COPM	Change score
1	4.2	7	2.8
2	4.2	5.6	1.4
3	– –	–	
4	5.2	7.2	2
5	5.2	7.4	2.2
6	5.6	6.6	1
7	4.4	6	0.6
8	3.5	5.5	2
9	3.6	5.8	2.2
10	3	7.4	4.4
11	2.2	4.4	2.2
12	4	6.5	2.5
13	3.6	7.4	3.8
14	4.6	6.4	1.8
15	3.6	4.4	0.8
16	5	6	1
17	1	5	4
18	4.3	5	0.7
Median score	4.2	6	2.0

2.2.4.3.3 Kinematic changes

The children's arm kinematics showed significant improvements in a number of movement parameters, including smoothness of movement, movement time and peak speed for both reaching and retrieve movements (Weightman et al., 2011).

2.2.4.4 Discussion

The feasibility of using the assistive computer games system in a home setting was satisfactorily demonstrated by this small study, with some useful suggestions from children for future development of the games system. There are promising indications that functional and kinematic benefits might result from use of the games system. However, there were a number of limitations with this study. Firstly, children did not withdraw from any therapy interventions that were in place, and without a control group the possibility that the children's ongoing therapy programmes were the cause of the improvements in activity limitation and arm kinematics cannot be ruled out. There was no correlation between individual children's improvements and the amount of time that they played the games. The amount of exercise undertaken was substantially less than the times described in studies to evaluate CIMT and bilateral therapy, suggesting that there may have been other mechanisms at work. Nevertheless, the aims of the study were achieved, and parents and children were positive about the potential of the games system.

Non-empirical observations suggest that some children showed reduced developmental disregard of the impaired arm, for example using the arm spontaneously for reaching and other simple tasks instead of performing inconvenient and uncomfortable posturing to bring the non-impaired arm into use.

A further unreported observation of potentially great importance in relation to time spent playing the games was parental engagement and support. In the feasibility study by Gordon et al. (2012) one child was withdrawn because the parents lost interest. In this study, the child who played the games system the most times and achieved the highest cumulative number of minutes played was strongly encouraged by their mother. She drew up a weekly chart and purchased stickers so that if the child played a certain amount of time per day they were rewarded with a sticker for that day. At the end of the week, a complete row of stickers earned the child a treat. This child also scored the third highest

improvement on the COPM outcome measure. This suggests that parents could have a more important role to play, providing additional stimulus and exercising some 'parental power', even when the children themselves are engaged and motivated to play with the games system.

The children reported that they would prefer to play the computer games with friends or siblings, suggesting that a collaborative or competitive element could have improved the usage of the games system.

The motors in the commercial joystick that was adapted for this study were not powerful enough to assist arm movements for children whose arm had more than a mild impairment, or to overcome the restriction to movement imposed by more severe muscle spasticity.

It was realised that the activity limitation outcome measures were not satisfactory for this study. The MABC is not validated for children with cerebral palsy, and had a large floor effect. The COPM was useful for identifying individual improvements, but the individualised nature of the goals means that it is difficult to compare across groups of individuals. However, it was fortunate that many of the goals in the 18 children were common to each other and allowed us to perform a statistical analysis using a non-parametric procedure (the Wilcoxon signed ranks test). Clearly, more diligent selection of outcome measures is essential, although other authors have highlighted the lack of appropriately-validated, responsive measures as an issue that needs attention (Eliasson et al., 2005, Hoare et al., 2007a, Hobart et al., 2007, Sakzewski et al., 2011a).

2.2.5 Conclusion

This feasibility study supports suggestions from case studies and other feasibility studies that potential exists for the use of assistive rehabilitation gaming technology to be installed in the homes of children with cerebral palsy. Feedback from the children suggests that collaborative and competitive games played with other children might potentially increase the amount of time spent playing the games system. The motors in commercially-available gaming joysticks do not provide enough assistance to children with a greater degree of impairment. There is a lack of responsive, validated measures of upper limb activity

available for the evaluation of changes in upper limb function of children with cerebral palsy.

2.3 Design, construction and feasibility study of school-based computer-assisted arm rehabilitation games system for children with cerebral palsy (NIHR-funded study, grant ID K005)

The outcomes of the G006 study outlined above suggest that children would prefer to play games in a social situation, rather than independently against a computer or themselves. It was also noted that more power would be necessary for the assistive mechanical support if the potential of robotic-assisted rehabilitation to include more severely-impaired children that was suggested by Fasoli et al. (2010) and Sandlund et al. (2009) was to be realised. It was therefore decided to objectively assess whether children preferred playing games by themselves or with friends, in both collaborative and competitive game play. To achieve this, a new device was constructed, designed and built with guidance and support from the user group of families of children with cerebral palsy.

Testing the robustness, safety and transportability of the device was planned as part of the study protocol. The study was funded by the National Institute for Health Research (NIHR) and adopted by the NIHR onto their portfolio (research network portfolio ID 6306). Favourable ethical opinion was obtained from Leeds (West) Research Ethics Committee (REC ref: 09/H1307/48) on 23rd June 2009.

2.3.1 Design of K005 new games and system: user group meetings

The children known to us from our previous research were invited to guide development and testing of the new games and device. These user group meetings took place at weekends and school holidays once or twice a year at the CRT Laboratory. The meetings opened with a brief presentation to describe the goals and aims and any progress since the last meeting.

At the initial meeting, children were part of small focus groups which were facilitated by researchers to find out which TV shows, characters and games were popular within the age group of our children (5 – 12 years old). Colouring pencils and paper were provided for children to contribute characters and ideas for games. Subsequent meetings provided

opportunities for children to play the prototype games, and review the developing games system. The use of focus groups and questionnaires to provide the opportunity for children to develop their own ideas and to give feedback was repeated. Activities such as these are illustrated in Figure 2-5 below.

Figure 2-5. User group meetings: testing of the games system; evaluating the hardware and the games; providing feedback through questionnaires and semi-structured discussions.

2.3.1.1 The games

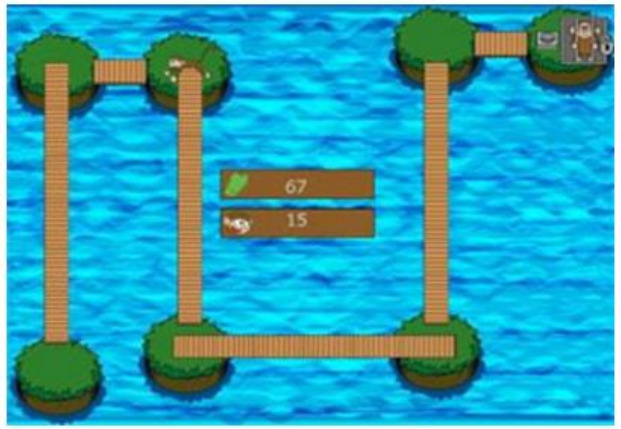
The new games were a mixture of collaborative games in which children played together using teamwork to achieve the objective, and competitive games in which children played against each other. It was also possible to play the games independently, against the

computer or themselves. The four games that the children conceptualised and helped to design are shown in Figure 2-6.

The river game. Children navigate the course of the river, collecting bananas for points. Penalties are incurred for beaching the logs.

The van game. The monkeys catch bananas as they swing through the trees. The monkeys fire the bananas at the van to release captured comrades by entering the chase vehicle, being careful not to lose points by being caught in the exhaust backfiring (the flame).

The puzzle game. Children take it in turns to cross bridges, press buttons and find keys to unlock doors. It is essential to work together to be able to release their comrade.



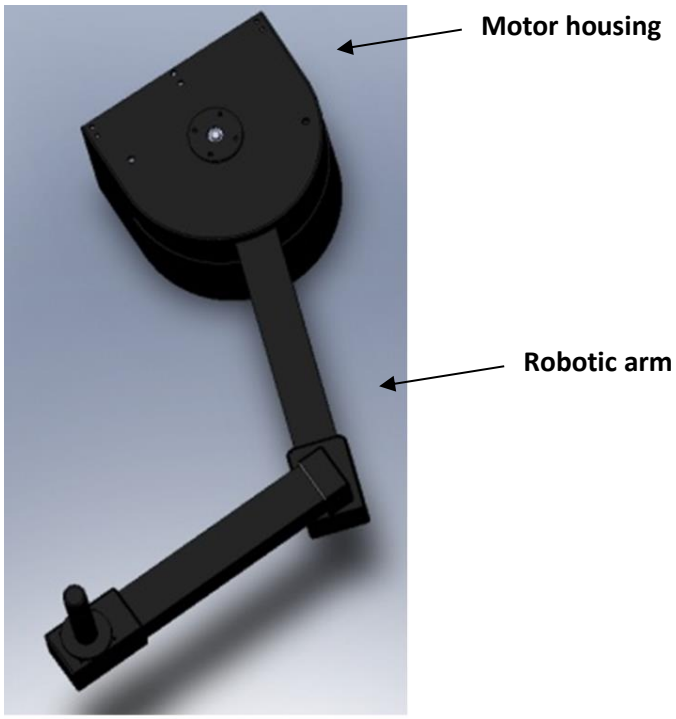
The chase game. Children follow the bridges to complete a route as fast as they can, against themselves or against their friends (each child taking a turn).

Figure 2-6. The four games designed by the children for the new assistive robotic games system.

2.3.1.2 The computer-assisted arm rehabilitation device (CAAR)

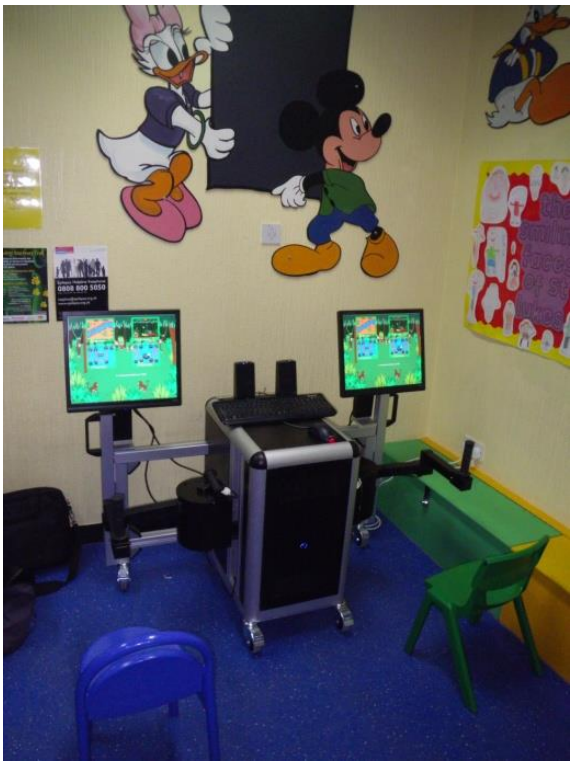
The new device was designed to withstand repeated use in a number of different geographical locations in a large area within northern England (West and North Yorkshire). Mobility was therefore also an important consideration. Most importantly, the device was built to European Union safety standards and included a number of precautionary fail-safe mechanisms to prevent the robotic arm causing any harm to the user in case of malfunction. This was particularly important given the large forces which could be delivered by the motors that powered the robotic arm; these were housed within the robotic arm itself (see Figure 2-7).

Figure 2-7. The robotic arm and motor housing.



As well as help from our user group the resulting device was tested in a busy paediatric outpatients clinic in a large regional hospital (see Figure 2-8 below), inviting children of all ages with varying disabilities and conditions and their parents and siblings to play the games system and give feedback using questionnaires and semi-structured interviews.

Figure 2-8. A dual user device deployed in a busy children's outpatients' waiting room at a large teaching hospital in the United Kingdom.



The objective of this exercise was to establish whether the robotic arm was sturdy enough to stand repeated and vigorous handling as well as test the device's mobility and general robustness to transportation and movement around a location to which it might be deployed. This stage of the study demonstrated that the games had the potential to engage children, that the device was mobile and that it was readily transportable by the research team (see Figure 2-9 below). Deploying the device required two members of the research team and a typical family estate car, and required removal of the computer screen.

Figure 2-9. Transporting the device using a family hatchback.

If necessary, the robotic arm could also be removed e.g. for ease of transport and for installation or entry to a deployment site if stairs or narrow doorways were an obstacle. Once within a building, however, the device was mobile because of the trolley on which it was mounted. It was anticipated that in future studies the device might be used within a child's home, so the device design had taken into account the typical size of doorways in UK dwellings. Therefore no difficulties were encountered deploying the device into school or hospital buildings, although some older school buildings required manual handling in the absence of lifts when the device was to be installed on an upper floor. Holt et al. (2013) describe the games system, its development and the games themselves in more detail.

2.3.1.3 Trunk restraint and seating

The decision was taken to not provide specialist seating or trunk restraint, the use of which suggests improved kinematics but not improved function of the upper limb (Schneiberg et al., 2010). Provision of these supportive elements is a resource-consuming process which could have serious delaying consequences for the study. It was also reasoned that if children recruited to the study required these aids then they would already have been equipped with them by their clinical support staff.

2.3.2 Feasibility study of the deployment of the CAAR games system in schools as a potential rehabilitation device

This stage of the study was to establish the feasibility of deploying the CAAR games system to English schools. The aim was to evaluate whether the games system had the potential

not only to engage children in the schools setting but to ascertain whether there was potential for the CAAR games system to supplement a rehabilitation programme during the school day. The primary aim of the study was to investigate whether children played the games system more if they used the system in collaborative mode (playing with school friends) or in independent mode (playing by themselves). This would establish an evidence base to direct future design of the most engaging and appropriate games for encouraging children to undertake therapeutic arm movements. Secondary aims of the study were: 1) to evaluate whether arm activity improved as a result of using the CAAR games system; 2) to evaluate whether arm kinematics improved as a result of using the CAAR games system; 3) to establish whether the system could be successfully integrated into school timetables; and 4) to carry out an initial validation study into a portable device for the kinematic assessment of the upper limbs of children with cerebral palsy in a non-laboratory setting. Holt et al. (2013) gives a full account of the success of the system's integration into the school routine.

2.3.2.1 Participants

Children eligible for the study were identified through local paediatricians and occupational therapist teams, but also included children from our user group. Inclusion criteria were children with cerebral palsy aged between five and twelve years old, who had upper limb activity limitation and cognitive ability to understand and play simple computer games. Twelve children with cerebral palsy were identified but one withdrew when their school (a secondary school) refused to participate because of the intensity of their curriculum. Eleven children in nine schools (eight primary schools and one secondary school) took part in this stage of the study (eight boys, three girls, all with unilateral impairment but for one child with bilateral involvement, aged from six to twelve years old (mean age nine years, SD one year eleven months)).

2.3.2.2 Study design

The study used a cross-over design (AB-BA), where A indicates dual-user mode and B indicates single-user mode. Children played in each mode for four weeks at a time, each separated by a minimum of three weeks 'wash-out' period. The wash-out period was by necessity timed to include school holidays. A period of games system maintenance was included between deployments, therefore each deployment of the games system took at least twelve weeks. Figure 2-10 shows the CAAR games system set up for each mode. The CAAR device was programmed to permit access to only one of the four games at a time,

inhibiting each game and enabling subsequent games every four days until the final four days of the deployment. For the final four days of the four week (20 school days) deployment, the children could select any of the four games to play. This option was designed to maintain interest in the games and to increase the likelihood that children would identify their favourite game through the amount of play time it received in the final four day period.

Figure 2-10. The CAAR games system set up for single user mode (left) and dual user mode (right) within two different schools.

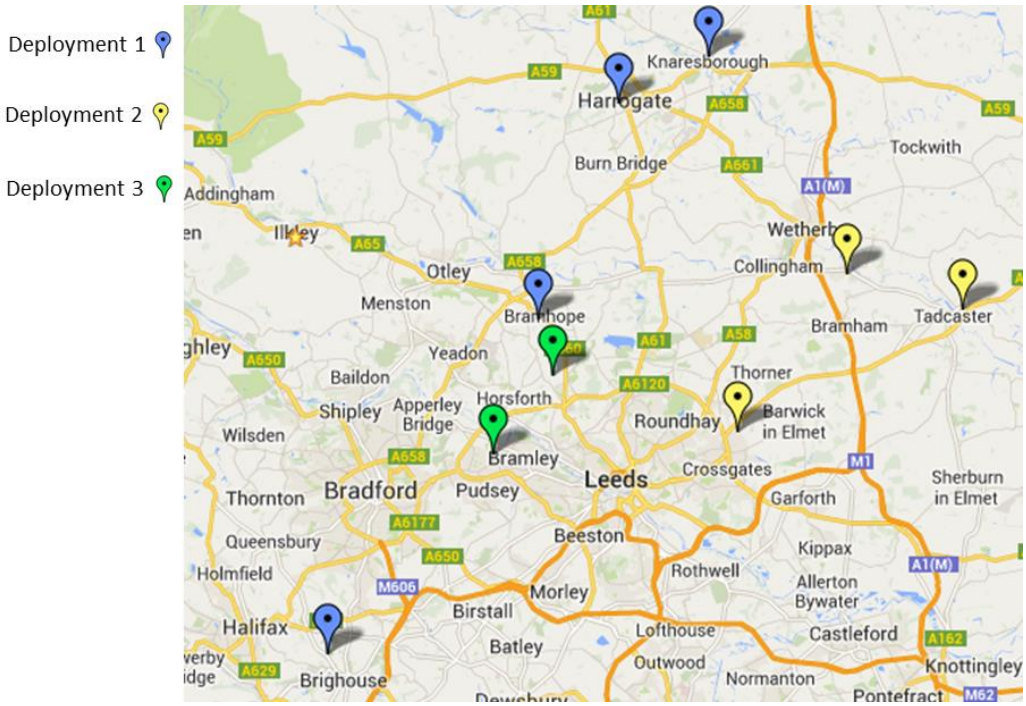
For the dual-user mode, children selected up to four non-disabled friends from their school with whom to play. Only one parent refused consent for their non-disabled child's participation, believing erroneously that the games were played during (instead of between) lessons.

2.3.2.3 Deployments

With four systems and eleven children, it was necessary to conduct the study through three stages of deployments. Children were randomly allocated to playing dual user (A) or single user (B) first, but the deployments were organised around geographical locations for convenience and efficiency of deployment, maintenance and collection visits (see Figure 2-11 below).

1. Deployment 1. Four games systems, used by five children in four schools.
2. Deployment 2. Three games systems, used by three children in three schools.
3. Deployment 3. Two games systems, used by three children in two schools.

Figure 2-11. Map showing general area of deployments.



2.3.2.4 Exercise regime

Other than request that children were allowed access to the system whenever it was appropriate within the school timetable, instructions to school staff were not prescriptive about use of the CAAR games system. It was suggested to staff responsible for supervision of the games system that children achieved thirty minutes of use a day, not necessarily in one games session. Any other rehabilitation programmes were to continue as detailed by the child's physiotherapist or occupational therapist.

2.3.2.5 Outcome measures

Measures to record changes in arm activity and arm movement were recorded at five time points:

1. Time point 1: two to four weeks before the child began using the games system (a Control assessment);
2. Time point 2: Baseline assessment, within three days of the child beginning the first four week deployment;
3. Time point 3: within three days of the child completing the first four week deployment;

4. Time point 4: within three days of the child beginning the second four week deployment;
5. Time point 5: Final assessment, within three days of the child completing the second four week deployment.

2.3.2.5.1 Amount of use in each mode

The primary aim of investigating which mode was most popular with children was evaluated by comparing the amount of time played in each mode (A versus B). The amount of actual game time was recorded by the CAAR device, so this was easily obtained at the end of each deployment. Children and supervising staff were also asked to complete questionnaires.

2.3.2.5.2 Changes in arm activity

Changes in arm activity were evaluated using the ABILHAND-kids and the COPM, which were described in paragraph 1.6.4.1 above.

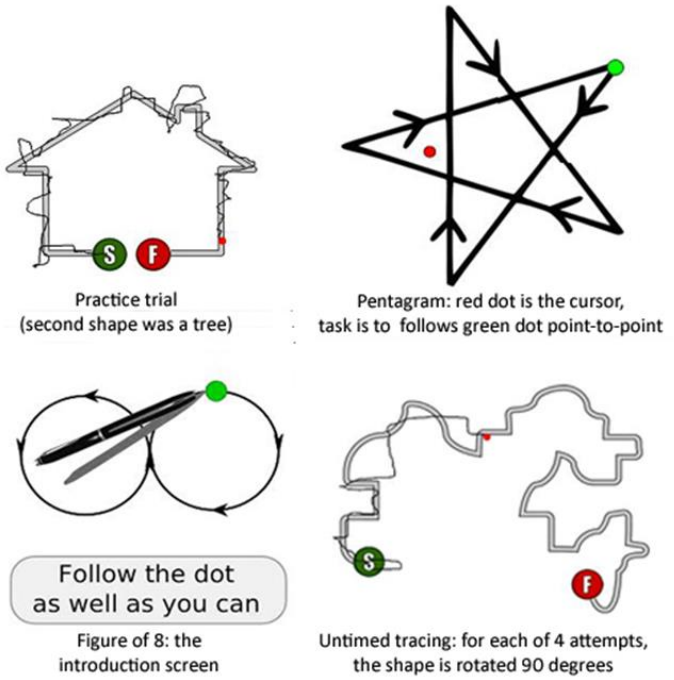
2.3.2.5.3 Arm kinematics

In the previous study, a CRT Laboratory based movement recording system was used to evaluate arm kinematics. For this study, it was unreasonable to ask the children and their families to visit the CRT Laboratory on five separate occasions. It was also anticipated that some families would not be able to attend the laboratory for all time points. Therefore all assessments were carried out at the children's homes. This presented no difficulty with the functional measures, which are designed and validated for use in clinics and residential locations, but it presented a problem for the evaluation of changes in arm kinematics. To address this problem a portable kinematic evaluation tool called the Clinical Kinematic Assessment Tool (CKAT) (Culmer et al., 2009) was adopted.

The CKAT was developed and tested for non-disabled adults using a tablet and a stylus. CKAT captures spatiotemporal movements of the upper limb as the user undertakes a number of tablet-based tasks using the stylus. These tasks attempt to mimic the type of paper-based hand-control assessments similar to those in the MABC. There is support for the development of robotic technology to assess motor changes in children with cerebral palsy (Frascarelli et al., 2009), so CKAT was adapted for children with cerebral palsy. This adaptation used a laptop with the adapted Microsoft Sidewinder joystick (used in the

home-based study described in section 2.2) as the user-laptop interface. Their tasks consisted of practice, speed, tracking and tracing tasks which were designed to elicit a variety of different movements. This adapted version of CKAT, called CPKAT, was able to record the movements of the onscreen sprite which essentially mimicked the children's temperospatial hand movement on the adapted joystick; it was therefore able to record accuracy, smoothness and speed of different trajectories in the horizontal plane. The tasks are shown in Figure 2-12 and described below. A sub-study to assess the feasibility of the CPKAT for use with children with cerebral palsy was designed and carried out as part of these assessments (Preston et al., 2014a). This sub-study is described in Chapter 3, sub-section 3.6 on page 196.

Figure 2-12. The CPKAT tasks.



Practice trial. Two practice sessions to familiarise the child with the joystick and control of the cursor movements consisting of tracing a simple house shape followed by a simple tree shape.

Pentagram: aiming task. The aiming task consisted of two attempts at a series of aiming movements around a Pentagram shape, guided by a target that moved with each successful aiming motion from point-to-point.

Figure of 8: tracking task. Four timed tracking tasks: the children track as closely as possible a target circle moving in a horizontally-positioned Figure of 8. The first two tasks are at a slow speed and the second two at a fast speed; each task lasts 31s. Speed of the task is

pre-determined and fixed. Children were asked to match the speed and position of the green circle.

Tracing task. Four untimed tracing tasks (identical shape, rotated 90° each time). There is no time limit, children were asked to take their time and to trace the shapes as accurately as possible.

By recording movement of the screen cursor CPKAT captures the movements of the child's hand and therefore allows some general conclusions to be drawn above the child's arm kinematics, These are defined in defined in Table 2-3 below.

Table 2-3. The parameters recorded by CPKAT to capture information about arm and hand movements.

TASK & OUTCOME PARAMETER	DESCRIPTION	REASONING
PENTAGRAM		
Path Length	The distance travelled during each point-to-point movement.	Indicator of overall performance.
Path Length Time	The time taken to travel the path length.	Indicator of hand speed.
Normalised Jerk index (NJ) (smoothness)	Measure of the smoothness and time taken for a discrete movement. A maximally smooth point-to-point movement of the Pentagram has a smoothness index of 7.75	Smooth movements are more energy efficient.
FIGURE OF 8		
Path Length	The distance travelled during each Figure of 8.	Indicator of overall performance.
Path Length Time	The time taken to travel the Figure of 8.	Indicator of overall performance.
Normalised Jerk index (NJ) (smoothness)	Measure of the smoothness and time of the tracking movement. A maximally smooth movement has a smoothness index of 7.75	Smooth movements are more energy efficient.
Path accuracy (RMS mean)	The position of the onscreen cursor controlled by the child via the joystick is monitored with reference to the position of the target moving along the Figure of 8 trajectories. RMS mean is a value of mean error.	Indicator of overall performance.

TRACING		
Path Length	The distance travelled during each shape.	Indicator of overall performance.
Path Length Time	The time taken to travel the shape's path length.	Indicator of overall performance.
Normalised Jerk index (NJ) (smoothness)	Measure of the smoothness and time taken for a discrete movement. A maximally smooth movement has a smoothness index of 7.75	Smooth movements are more energy efficient.
Path Accuracy	The sprite trajectory compared against the shape's reference trajectory.	Indicator of overall performance.
TPA (Time/Path Accuracy)	A product of Path Accuracy and Path Length Time; TPA allows comparison of children who sacrificed speed for accuracy and vice versa.	Indicator of overall performance.

TPA: Time/Path Accuracy; NJ: Normalised Jerk; RMS: root mean square

During the CPKAT assessment, each child sat in a standardised position at the dining or kitchen table as illustrated in Figure 2-13 below. All children but one sat on a typical dining room or kitchen chair with feet supported on a firm surface, the remaining child sat in specialist seating with trunk and foot support, as provided by the local paediatric services.

Figure 2-13. Five year old using CPKAT device to evaluate arm kinematics (child not a study participant).

All distractions were removed or minimised (e.g. siblings exiled, television off). Children wore hand orthotics (e.g. hand splints) when using CPKAT if orthotics were usually worn

for daily activities. The CPKAT assessor sat alongside the child to explain each task and to offer encouragement. The adapted joystick was placed between the child and the laptop, as close to each as possible while allowing for a full range of movement of the joystick. The joystick was used unpowered (no assistance or feedback from motors).

2.3.2.5.4 Statistical analysis

To determine whether children played the games system more in single-user mode (playing by themselves) or in dual-user mode (playing with school friends), the non-parametric Wilcoxon Signed Ranks Test was used, with alpha set at 5%.

To determine whether any changes in activity took place that could be attributed to the games system the non-parametric Wilcoxon Signed Ranks Test was used with alpha set at 5%. Activity limitation outcome measure scores were analysed from baseline to final assessment (time point 2 – time point 5).

Kinematic performance across each deployment and from baseline to final assessment was assessed using the non-parametric Friedman's ANOVA to determine whether there were differences between measures at each time point, with post hoc tests performed if appropriate using the Wilcoxon Signed Ranks Test with alpha set at 5%. Statistical analysis was carried out using PASW Statistics 18 (Release 18.0.3).

2.3.2.5.5 Results

2.3.2.5.5.1 Use of the system in dual and single user mode

Comparison of usage in dual-user and single-user mode was not statistically different, with children playing each mode almost equally. The eleven children played the games system for a total of 132 days in single use mode (median days used per child was 13 days) and 121 days in dual use mode (median days used per child was 12 days). However, part of this difference can be accounted for by two children who did not play the game at all in dual-user mode. The median daily use was 9.27 minutes in single use mode and 11.2 minutes in dual use mode. Though this difference may appear to show a preference for dual-user exercise it was not statistically significant (p = 0.214, based on children's median daily use). However, the children indicated a clear preference for dual-user mode on the questionnaires which they completed at the end of their participation in the study.

2.3.2.5.5.2 Arm activity changes

The outcome scores of the two measures used to evaluate changes in arm activity showed no significant difference before and after use of the games system. Table 2-4 below shows arm activity changes on both the ABILHAND-kids and the COPM for all eleven children from base line (time point 2) to the final assessment (time point 5) i.e. after approximately eight weeks of game play with a three to six week washout period after four weeks. On the primary outcome measure (the ABILHAND-kids), five children showed activity improvement, two showed deterioration in arm activity and four showed no change. On the COPM, two children showed arm activity improvement and nine showed no change. Differences between the baseline measures and final measures were not significant (ABILHAND-kids, $p = 0.424$; COPM, $p = 0.484$).

Table 2-4. Outcome data for usage and arm activity changes.

	Number of days played (days)		Max time played in a day (mins)	Total time used per child (mins)		Total use (mins)	Median daily use (mins)	Median per day (mins)	Change in ABILHAND-kids score (logits)	Change in COPM score (out of 10)
Participant	Single use	Dual use		Single use	Dual use	Single and dual use combined (mins)	Single use	Dual use	Change of more than 0.45 indicates functional difference	Change of 2.0 or more indicates functional difference
1	10	18	11.24	47.27	123.26	170.53	4.55	6.37	-0.169	1
2	4	12	8.79	17.28	104.50	121.79	3.57	8.98	0.657	1.6
3	15	16	24.43	239.38	262.77	502.15	14.55	15.65	-0.645	-1.8
4	15	19	39.69	298.83	297.95	596.78	18.79	14.07	-0.409	-0.2
5	13	9	10.18	98.20	80.12	178.31	8.38	10.18	0.632	0.8
6	16	12	21.41	175.74	133.27	309.01	11.71	10.85	0.875	2.2
7	14	9	26.94	182.24	65.52	247.76	12.52	7.12	-0.844	-0.8
8	10	15	29.31	112.96	286.36	399.32	11.34	20.36	0.534	1.8
9	7		6.54	21.45		21.45	3.23		1.473	3.2
10	12		23.60	149.88		149.88	11.06		0.305	0.66
11	16	11	20.36	110.61	157.22	267.83	5.67	15.24	-0.439	0.8
Median	**13 days**	**12 days**	**TOTAL:**	**1453.83**	**1510.97**		**11.06**	**10.85**	**0.305**	**0.8**
TOTAL DAYS	132 days	121 days							No difference, p = 0.424	No difference, p = 0.484

2.3.2.5.5.3 Arm kinematics

The CPKAT results showed significant changes in the children's arm kinematics. Children showed an improvement in hand speed and in smoothness and accuracy of movement (see Table 2-5).

Table 2-5. Results of CPKAT, showing improvements in speed, smoothness of movement and accuracy.

TASK **Outcome Parameter**	**RESULT**
PENTAGRAM (aiming movements)	
Path Length (mm)	No difference between groups, p = 0. 445 (Friedman's ANOVA).
Path Length Time (seconds)	*Difference detected time point 2 to time point 3, p = 0.028 No difference baseline to final assessment (p = 0.508)
Normalised Jerk index (NJ) (smoothness) no units	*Difference detected time point 2 to time point 3, p = 0.005 No difference baseline to final assessment (p = 0.241)
FIGURE OF 8 (tracking task)	
Path Length (mm)	*Difference detected time point 4 to time point 5, p = 0.022 No difference baseline to final assessment (p = 0.241)
Path Length Time (seconds)	No difference between groups, p= 0.222 (Friedman's ANOVA)
Normalised Jerk index (NJ) (smoothness) no units	*Difference detected time point 4 to time point 5, p = 0.047 No difference baseline to final assessment (p = 0.799)
Path accuracy (RMS mean) no units	*Difference detected time point 4 to time point 5, p = 0.037 No difference baseline to final assessment (p = 0.203)
TRACING	
Path Length (mm)	*Difference detected, time point 4 to time point 5, p = 0.028 No difference baseline to final assessment (p = 0.203)

Path Length Time (seconds)	No difference between groups,, p = 0.398 (Friedman's ANOVA)
Normalised Jerk index (NJ) (smoothness) no units	No difference detected, p = 0.398 (Friedman's ANOVA)
TPA no units	*Difference detected time point 2 to time point 3, p= 0.022 *Difference detected Baseline to Final Assessment (p = 0.007)

TPA: Time/Path Accuracy; NJ: Normalised Jerk; RMS: root mean square

2.3.2.5.6 Discussion

No difference was found in the amount of times that children played the games system in each mode, with children playing each mode almost equally. The reason that playing times between single and dual-user did not differ was due to school timetabling. Within each school, each mode was restricted to similar amounts of available playing time thus preference for game play in either mode was unable to influence the total time played. However, qualitative reports from the children indicated that they strongly favoured the collaborative mode (playing with their friends in non-competitive games). Between schools, playing times of the games varied considerably, as shown in Table 2-4 on page 76: child 3 and child 4, both attending the same school, exercised on over 30 days of the 40 days the system was available to them in the trial period while two other children (child 9 and child 10, both attending the same school) played for less than 15 days. In fact, child 3 and child 4 each used the games device more than any other child, and this is notable for two reasons: firstly, they were at a secondary school, and another secondary school had refused on the grounds that secondary schools were far too busy with National Curriculum activities to support research studies like this; and secondly, the Special Educational Needs teacher was extremely supportive of the children, their requirements and of any strategy with the potential to benefit them, including in his strong and unwavering support for their participation in this study.

This final point suggests that adequate use of the CAAR device, and perhaps participation with any rehabilitation device or programme, in schools depends on the school's commitment and support. Schools were approached only because the child and parents had expressed their eagerness to take part, and the schools were therefore obliged to

support these children (though not in the case of the secondary school who refused to participate). Non-empirical observations suggest that the amount of time that the CAAR device was played by the children was reflected by the level of support and commitment displayed by each school.

Further analysis of the data shows that while in some schools the average use was over 19 minutes per day in others it was as little as five minutes. This is less than the initial target of 30 minutes and indicates how difficult it is to take significant amounts of time out of the school day in a busy school's curricula. This is particularly noticeable at certain points in the calendar. For example, child 9 and child 10 did not play the games at all during the second deployment (when they should have played dual-user with friends) because of the intensity of the opening weeks of the school year following their transition to a higher school year, and child 11's opportunity to play during the dual-user deployment was substantially reduced because National Curriculum assessments were underway in his year group. Other traditional events in the school calendar also impacted on usage e.g. the Christmas Nativity play. One further problem prevented a child from undertaking exercise with their affected upper limb for two weeks because the supervising adult had an apparent misunderstanding and refused to let the child use their affected arm to operate the robotic arm.

No changes were found in upper limb activity before and after use of the games system. This is not to say that the contribution made by the games would not be of benefit when used in combination with a home-based system, with a concentrated burst of therapist-facilitated rehabilitation or following surgery or medical intervention e.g. botulinum toxin treatment for spasticity. This finding that arm activity did not significantly improve may be due to the poor responsiveness of the measures used. More likely, the limited time that the children played the games is unlikely to have made any functional impact. However, while the change in scores is not significant, there is a noticeable trend towards an inverse correlation between usage and changes in functional ability. For example the two children that used the games system the most (child 3 and child 4) showed a *clinically* significant deterioration in activity performance over the study period based on both the ABILHAND-kids and COPM scores, while child 9 showed the largest clinical improvement even though they played the game system the least amount of time.

On the other hand there were statistically significant improvements in children's arm kinematics. Improvements were seen in the Pentagram task for speed and smoothness of hand movements, in the Tracking task for path length, smoothness and accuracy and in the Tracing task for path length and Time/Path Accuracy (TPA). There was no improvement in the Pentagram's Path Length, Figure of 8's Path Length Time and Tracing's speed and smoothness. However, the Tracing task is not about speed and children were encouraged to take their time and concentrate on accuracy; this is captured by TPA, which showed a significant improvement. The act of taking their time on this task also caused the movements to be cautious and broken, so smoothness was perhaps a poor parameter to measure for this task. The Pentagram showed no change in Path Length which may indicate that the children were accurate at following the set pattern, but gained arm speed and smoothness. The Tracking task showed no difference in Path Length Time, which is unsurprising since the speed of the object to be tracked is pre-determined, fixed and constant. Tracking also indicated improvements in accuracy (also indicated by improved path length) and smoothness, and Path Length and TPA (signifying improvements in accuracy and time taken) improved for the Tracing task.

Nevertheless, the observed kinematic improvements after less than 15 minutes activity per day suggest the potential for activity improvements if more time is spent playing the games system, perhaps in combination with a home-based system so that the amount of therapeutic activity achieved in other trials that produced benefits in functional activity is achieved e.g. 75 minutes per week (Knox and Evans, 2002). The kinematic improvement suggests an additional potential benefit for the use of kinematic analysis, perhaps providing evidence that a greater intensity of the intervention might go on to produce functional benefits.

A number of potential issues with use of CPKAT were identified. The number of CPKAT tests at each assessment proved difficult. Some children clearly tired of the assessments towards the end and strayed from the task in hand, for example demonstrating to the assessor on one occasion how fast they could make the cursor go in circles, thus rendering that task's measurements very inaccurate. In future, engagement with these CPKAT assessment tasks might be increased by incorporating them within a game. Some of the

tasks capture data that is also captured in other tasks, so limiting the number of tasks might prevent loss of interest and concentration.

Non-disabled children were able to play on the games system without interference in their academic programme, supporting their friends who have cerebral palsy in a potentially useful rehabilitative activity that also increases the social contact of the child with cerebral palsy in their peer group.

2.3.2.5.7 Conclusion

Computer-Assisted Arm Rehabilitation systems can be successfully installed in schools, but their daily use depends on the academic year and the pressures of the current school schedule e.g. transition pressures early in the new school year or National Curriculum assessments. Enthusiastic support of school staff also appears to be an important element. Children prefer to play against or with their friends rather than on their own, increasing the value of school-based deployments, however the time available for using the system (and receiving adequate rehabilitation) is limited.

Arm activity showed no significant changes in the limited time that children played the games system, but there were significant improvements in arm kinematics. This suggests that twelve minutes per day is insufficient to achieve benefits in arm activity limitation but does improve arm kinematics.

Assistive gaming technology can successfully be installed in children's homes and schools, with potentially functional and kinematic benefits. Although encouraging a sufficient quantity of time on the gaming technology might still present a challenge, use of the gaming technology might complement an established rehabilitation programme particularly in the small window of opportunity following botulinum toxin when targeted and intensive therapy is crucial.

2.4 Summary of the development and use of assistive robotic gaming technology

There has been an increase in the numbers of studies exploring the use of virtual reality and high quality video games e.g. Nintendo Wii®™ in the rehabilitation of disabled adults and children. The papers generally describe feasibility studies, case studies and studies of

limited quality but the results suggest some potential of the technology to engage children in intensive, repetitive therapeutic movement exercises.

Commercial games and hardware are not designed to provide targeted therapeutic exercise or movements, or assistance to children with impaired movement. Adapted hardware appears to lack the robustness and power to assist passive movement of a weak upper limb or to overcome the resistance offered by spasticity. This suggests that the best option for rehabilitation researchers is to develop their own assistive technology and programmes, but this is expensive in terms of time and finances.

Our own feasibility studies carried out with a number of children with cerebral palsy at home and at school further demonstrated the potential for this technology, with indications that there are functional and kinematic benefits to be gained. Children reported enjoying playing games with assistive technology when installed at home but derived greater pleasure when they were able to play with friends in either competitive games or games that require collaboration and team work to achieve the game's objective. This suggests that the interactive support of parents and school staff in rehabilitation programmes is essential if programme adherence and sufficient therapeutic time on the games system is to be achieved. However, it is likely that the amount of time available to play the games system in schools is too restricted by the school timetable to achieve any functional benefits, although it may still be enough to support a targeted rehabilitation programme following surgery or botulinum toxin treatment.

Botulinum toxin treatment is most effective when used as an adjunctive therapy in support of targeted rehabilitation programmes (Hoare et al., 2013, Hoare et al., 2010). Not all therapy teams have the resources to implement intensive and appropriately-targeted programmes to take advantage of the benefits of botulinum within the window of opportunity provided by this treatment (one team which supported one of our studies, discussed later, was down to two staff members and was not undertaking face-to-face contacts). This suggests that the use of assistive robotic technology would be useful to support rehabilitation programmes following botulinum toxin treatment.

According to several investigators exploring the potential effects on function of using assistive robotic technology, outcome measures for evaluating upper limb activity limitation in research are non-responsive and lack satisfactory psychometric evaluation e.g. validation for children with cerebral palsy (Eliasson et al., 2005, Hoare et al., 2007a, Sakzewski et al., 2009). Identifying an outcome score that can be used in parametric statistical calculations, that is responsive and appropriate for use with children with cerebral palsy is challenging but essential, if the potential of this technology – as well as that of other experimental interventions - is to be fully and properly evaluated.

It was therefore proposed to carry out a randomised controlled trial (RCT) of a gaming device which included assistive robotic technology used in support of botulinum toxin treatment follow up. The aim of the RCT was to investigate whether use of the device enhanced the functional benefits to the upper limb of children with cerebral palsy following botulinum toxin treatment. The first stage of the RCT will establish an appropriate measure for measuring activity limitation of children with cerebral palsy.

3 Measuring activity limitation and kinematics in children with cerebral palsy

"In physical science the first essential step in the direction of learning any subject is to find principles of numerical reckoning and practicable methods for measuring some quality connected with it. I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of Science, whatever the matter may be."

~ William Thomson, 1st Baron Kelvin (1824 – 1907)

3.1 Introduction

The measuring of certain common variables e.g. weight, length or temperature is generally a straightforward procedure. These variables are singular, discrete entities with which there are appropriate tools to carry out the measurement – weighing scales for weight, a ruler for length or a thermometer for temperature. We select a different weighing scale for measuring a small quantity of butter for a recipe to that for evaluating the impact of Christmas excess on body mass, while a thermometer used to measure the temperature of an oven being used to roast a chicken would be useless for measuring a patient's body temperature. We give little thought to these common tools when using them and we trust and rely on the numbers associated with the measurement even when our health and welfare depend on them. We trust that the particular tool of measurement in each circumstance is valid i.e. there is good evidence that the scale adequately measures the variable for which it is intended to measure (Bond, 2003, Bond and Fox, 2001, Hobart et al., 1996); we trust that the measurements it issues are reliable i.e. that the numbers produced by a scale to represent the quantity of the variable are consistent and reproducible every time, on every occasion and for each different person that uses it (Field, 2009, Hobart et al., 1996). Responsiveness is another important quality of measurement. A change of a degree Celsius in body temperature of a hypothermic or febrile child can be life-threatening so it is essential that the thermometer can capture this change.

Besides reliability, validity and responsiveness a measure should be unidimensional (Horton et al., 2013) i.e. scales should actually measure the characteristic that they say they measure, and not a different characteristic altogether (Hobart et al., 2007). This may seem obvious. However, for example, the MAUULF purports to measure unilateral upper limb activity limitation, and has been used for this purpose in clinical trials (Hoare et al., 2010) yet a large proportion of its items relate to ICF-CY Body Function categories (Hoare et al., 2011). The problem with this is that, for example, improvements in wrist extension and elbow extension will potentially indicate a clinically and statistically significant reduction in activity limitation when no such functional benefits have actually occurred. An example of a health measure which uses two characteristics is the Body Mass Index (BMI), a health indicator which uses height and weight to generate a single numerical figure as a measure of fatness (Garrow and Webster, 1985). A healthy BMI falls between the BMI values of 18 and 25, but professional rugby players – renowned for their athleticism – are classed as morbidly obese with a BMI far higher than 25 (King et al., 2005).

Other variables are not so manifest or accessible for direct measurement e.g. pain, anxiety, depression, intelligence. Methods of measuring these human characteristics include the use of questionnaires. Typically, questionnaires include a number of questions or statements known as items; each item has a number of response categories, and the respondent endorses the response that most closely indicates his 'level' of the variable that the item is measuring. It is essential that such measures are as unidimensional, valid, reliable and responsive as the thermometer or weighing scales described earlier.

A number of investigators carrying out research into functional benefits of experimental interventions have suggested that the measures they used to capture changes in activity lack appropriate validity for children with cerebral palsy, lack responsiveness to detect clinical or statistical change or are inappropriate for parametric statistical analysis (Hoare et al., 2010, Palsbo and Hood-Szivek, 2012, Qiu et al., 2009, Sandlund et al., 2009, Grimby et al., 2012, Hobart et al., 2007, Tennant, 2007). Many of these measures produce ordinal-level outcome scores. Hobart et al. (2007) present compelling arguments that only interval-level data is adequate for research, and argue eloquently that appropriately validated, responsive, psychometrically sound measures of upper limb activity that

produce a linear outcome score are essential for researchers and clinicians alike (Hobart et al., 2007).

These are serious obstacles to progress in establishing evidence-based interventions for rehabilitation programmes for children with disability, both in clinical practice and in research. Using poorly validated, unreliable, unresponsive or otherwise untrustworthy outcome measures, even with an otherwise rigorous methodology in a clinical trial or the most diligent adherence to a carefully planned and personally-tailored rehabilitation programme, casts doubt on any suggestions of resulting benefits.

It is therefore essential to identify a reliable, validated and responsive measure to evaluate the results of research and clinical intervention for improving upper limb activity limitation in children with cerebral palsy.

3.2 Introduction to psychometrics

In order to identify an appropriate outcome measure for evaluating changes in upper limb activity of children with cerebral palsy it is necessary to investigate the quality and level of psychometric testing that they have undergone to evaluate their own scientific standards of measurement. In order to evaluate the strengths of each measure, it is helpful to understand the psychometric techniques which have been used to develop and test these measures.

Traditionally, the reliability and validity of measures have been evaluated using Classical Test Theory (CTT; DeVellis, 2006). According to Hobart et al. (2007), CTT is based on an unsound scientific basis because it is untestable, and is the cause of many of the psychometric weaknesses outlined above. Furthermore, CTT tends to produce measures with many items (DeVellis, 2006); this can artificially raise the coefficient score that indicates acceptable reliability (Tennant and Conaghan, 2007). Many of these items are similar (DeVellis, 2006); this can lead to response dependency (Tennant and Conaghan, 2007) which can artificially raise outcome scores.

Modern psychometric techniques such as the Rasch model offer a number of advantages over CTT. The Rasch model is a mathematical measurement model based on the fundamental scientific principles of measurement (Tennant and Conaghan, 2007). A dataset formed from responses to questionnaire items are evaluated for their fit to the Rasch model to determine whether the questionnaire meets acceptable standards required for scientific standards of measurement. Familiarity with these approaches to psychometric testing is therefore helpful if one is to attempt an evaluation of the essential properties and psychometric standards of the measures.

3.2.1 An introduction to Classical Test Theory (CTT)

The basis for CTT is given by the formula

$$OS = TS + e$$

in which OS is the observed score, TS is the true score and e is the error associated with the observation or response to an item or question which seeks to measure the variable of interest (DeVellis, 2006). Over a number of items or observations, the error is expected to be random and therefore its mean error across items is zero (DeVellis, 2006). Item reliability and validity is evaluated on assumptions based on this principle (Hobart et al., 2007).

Item reliability assumes that a good item produces an observed score that is closely related to the true score, and that this observed score varies across individuals (because the variable of interest varies across individuals) and varies across time in one individual e.g. after an intervention (DeVellis, 2006). There is an implication that both true and observed scores covary, and that a measurement of the strength of this covariance will provide an indication of the accuracy of the observed score (DeVellis, 2006). The measurement of the strength of this covariance, and thus the reliability of the measure, is called the correlation coefficient, which if squared gives the proportion of variance shared between the true and observed scores. However, the true score is unobservable, inseparable as it is from the error with which it is associated. CTT holds that a number of assumptions combine to address this problem (DeVellis, 2006). Firstly, the items must be strictly parallel, which is an assumption that comes with its own conditions e.g. items' covariance with true scores must be equal across items and each item's error must be independent both of the true score and of every other item's error (DeVellis, 2006; page S51). This condition is itself

unprovable, and therefore CTT is a theory that cannot be “tested, verified, or - more importantly - falsified in any dataset” (Hobart et al., 2007, page 1098). Hobart et al. (2007) asserts that these are the reasons for the problems with measures described above. For example, because of the unknown qualities and quantities associated with the observed score, only the ordinal raw scores can be used for any statistical operations (non-parametric statistics), and the width of the confidence intervals for individual scores (Hobart et al., 2007) makes conclusions based on change scores between individuals very ambiguous. Also, DeVellis (2006) notes that ordinal scales show varying degrees of responsiveness along the scale from one extreme to the other. Hobart et al. (2007) suggests that this non-linear relationship can vary by a factor of up to 15 across the range of the scale. DeVellis (2006) notes that there are other psychometric issues with CTT, such as inability to detect response bias (differential item functioning (DIF)) and that the scales produced are sample dependent. DIF exists when a sub-group within a population endorses responses to items differently to other sub-groups, therefore obtaining a different score even though they have the same amount of the variable being measured. An example of this might be pain or the level of discomfort due to a common cold when measured on a scale given to a group that includes both men and women: one might find that, *despite the same level of discomfort,* women might endorse items to obtain a lower score for each variable than men.

DeVellis (2006) suggests advantages of CTT include the familiarity and easily understood concept of the reliability coefficient, and that methods of developing measures using CTT are readily available and easily managed. DeVellis (2006) further suggests that the nature of CTT-developed measures supports the use of less-appropriate items that contribute less satisfactorily to measurement, stating that “adding items can offset this problem and theoretically, just about any desired level of reliability can be achieved” (page S57). But these advantages do not counter the problems discussed above, and in fact adding items to improve reliability can add problems e.g. local dependency, which causes artificially increased scores.

The requirement in scientific and medical practice for high standards of measurement defines an outcome measure that is valid and reliable, invariant across samples, unidimensional, is free of locally dependent items and is free of DIF. The Rasch model is a

mathematical model which offers the opportunity to perform parametric statistical analysis of item responses, thus allowing an objective evaluation of the items' psychometric properties. Each item can then be modified to address the psychometric deficiencies or, if necessary, deleted altogether. Furthermore, because the Rasch model is based on fundamental scientific principles of measurement, it models the expected behaviour of the questionnaire response dataset for the questionnaire to achieve linear (interval-level) outcome scores (Tennant and Conaghan, 2007).

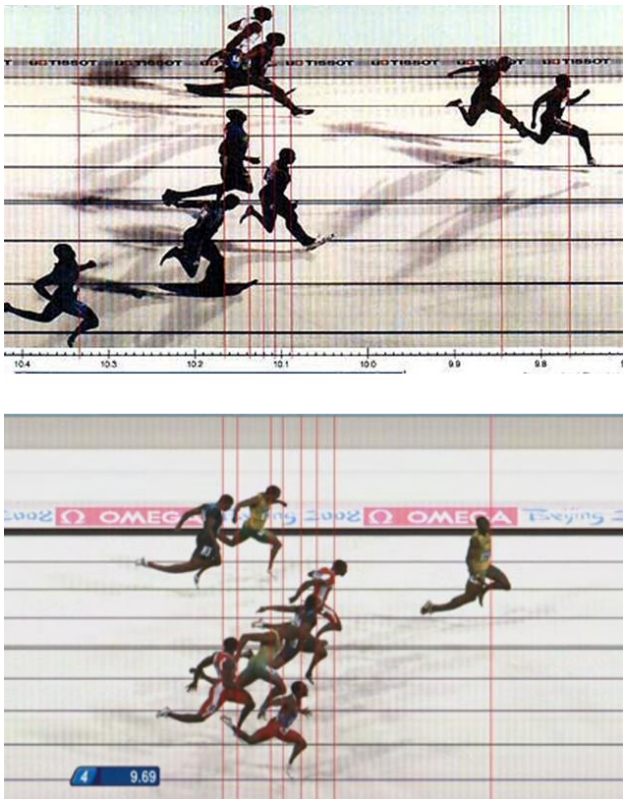
3.2.2 The problem with ordinal outcome scores

Many measures described as primary and secondary outcome measures for research studies produce ordinal-level outcome scores (Hoare et al., 2007a, Hoare et al., 2010, Sakzewski et al., 2009). The wide confidence intervals associated with ordinal outcome scores make these measures unsuitable for assessing clinical change in individuals nor are they appropriate for parametric statistical operations such as addition, subtraction, division or multiplication (Hobart et al., 2007). It has been suggested that the inappropriate use of raw scores as an outcome measure is damaging the research by health investigators to advance their areas of speciality into the realms of science (Tennant et al., 2004).

The recognition that ordinal data is misused, supports misinferences and is unhelpful to clinicians and researchers is not a recent advance (Merbitz et al., 1989, Stucki et al., 1996). The main problem with the use of ordinal data stems from its very nature, that of grouping similar levels of the variable being measured into different categories that can be ordered by size. The dividing lines between these groups are arbitrary (Merbitz et al., 1989) and each group is not necessarily of identical width or quantity, even when that quantity can be more accurately measured e.g. undergraduate degree classifications (Field, 2009). Often, the different groups are given consecutively-numbered identifiers, and this potentially causes the main problem: that of treating each category as being numerically equal (Merbitz et al., 1989). This is not a property of ordinal data. For example, the GMFCS described in Table 1-6 on page 10 is an ordinal classification system with five categories of ascending mobility limitation (Eliasson et al., 2006). Categories, or levels of mobility limitation, are numbered from I – V. The ordinal nature of the GMFCS means that there is increasing mobility limitation as the levels increase in numerical value i.e. Level I describes less mobility limitation than Level II. However, a change from Level I to Level II does not necessarily represent an equal change of mobility as a change from Level II to Level III.

Ordinal data can influence health and research outcomes with this misinference (Merbitz et al., 1989). If a child who was positioned near the top of Level II received an intervention that caused a small improvement in mobility – just enough to take them into Level I, this would indicate success for the intervention; however, the results of an intervention that caused a child's mobility to progress from the bottom of Level II to the top of Level II would indicate failure of the intervention despite potentially greater improvement in mobility. A clearer example of this problem is illustrated in Figure 3-1 below. This shows the photo finish of two 100 metres sprint events. For the athletes in third place to achieve a silver medal, one of them will have to work somewhat harder to make this transition to the higher ordinal category; furthermore, even if they each run faster by a second, equating possibly to a 10% improvement, they might still not achieve the improvement in ordinal outcome measures if the second place athlete also improves. The impression given by First, Second, Third is of equal intervals, and ordinal data is commonly treated this way (Merbitz et al., 1989).

Figure 3-1. Example of how ordinal data gives less information than interval level data.



It is these unknown and varying distances between and within categories that make ordinal data inappropriate and inaccurate for statistical operations of subtraction, addition,

division and multiplication (Merbitz et al., 1989, Svensson, 2001). Yet ordinal outcomes are frequently used in parametric statistical analysis, or presented as means and standard deviations (e.g. Wallen et al., 2007). The outcome score of the COPM is not only ordinal but is generated by addition and division of ordinal scores to create a mean score (Law et al., 2005).

Hobart et al. (2007) draw attention to the possibility that the confidence interval of an ordinal outcome score for individuals can occupy 30% of the available range of the scale, suggesting that a remarkable change in scores will be necessary before any significance can be attached to it. This has potential implications for the change in COPM scores that is regarded as clinically significant – a change of 2 out of 10. Hobart et al. (2007) further suggest that that ordinal scales are by their nature likely to be unresponsive. This could result in type-II errors. Arguments against more responsive scales include the likelihood of type-I errors (Hobart et al., 2007) but this is easily avoided by establishing a change score for the measure which indicates clinically significant change. Common sense suggests that for a measure to detect a clinically significant change in the characteristic being measured, its sensitivity and scale calibrations must be greater than that of the minimally significant change.

If ordinal-level outcome scores are unacceptable for health research, a psychometric model for establishing the fundamental properties for scientific measurement and interval-level measurement must be used. The Rasch model provides the opportunity to develop measures which meet these properties and to test for them in existing measures.

3.2.3 An introduction to the Rasch model

The Rasch model offers a number of advantages over CTT, including the capacity to determine whether the measure undergoing psychometric evaluation meets the fundamental requirement for producing interval-level outcome scores (Tennant and Conaghan, 2007). The underlying principle for the Rasch model is that the probability of a person endorsing, or 'passing', an item is influenced by both the difficulty of the item and the ability of the person (Tennant and Conaghan, 2007). That is, the hardest item on a measure will be endorsed or passed only by those with the greatest ability of the trait

being measured. Endorsing or passing an item illustrates a specific 'quantity' of the trait being measured; therefore it is probable that all easier items will be passed or endorsed by that person, and also probable that items of more difficulty will not be endorsed. The Rasch model therefore resembles a *probabilistic* form of Guttman scaling (Tennant and Conaghan, 2007).

Guttman scaling usually forms a deterministic pattern, as shown in the hypothetical example of a mobility measure in Table 3-1 below. Table 3-1 shows a sample of the population, each with different levels of limited mobility. Their mobility (ability to mobilise) is given by their capacity to achieve the items: '1' for indicating that they have the ability to achieve that item, '0' to indicate that they are unable to achieve that item. Item difficulty and person ability are arranged in ascending order.

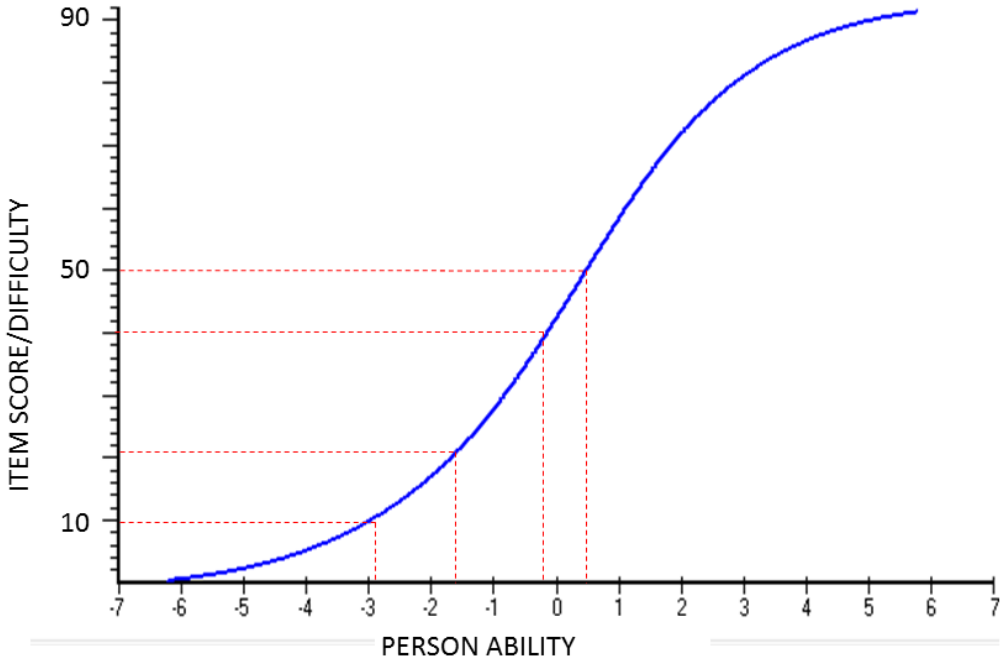
Table 3-1. A Guttman scaling pattern.

	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Person 7
Walk several miles for leisure	0	0	0	0	0	0	1
Walk 1500 metres (around local town)	0	0	0	0	0	1	1
Walk 500 metres (to shops, friend's house)	0	0	0	0	1	1	1
Walk 100 metres (to furthest end of road)	0	0	0	1	1	1	1
Walk 20 metres (upstairs to bedroom)	0	0	1	1	1	1	1
Walk 10 metres (to the dustbin)	0	1	1	1	1	1	1
Walk 5 metres (to next room e.g. kitchen)	1	1	1	1	1	1	1

In Table 3-1, Person 4 is able to walk 100 metres (to the end of their road). It therefore follows that they are able to walk all distances less than this. Equally, they have been unable to endorse the item 'Walk 500 metres (to shops or a friend's house)', and it follows therefore that they are unable to walk distances in excess of 500 metres.

The Rasch model is less strict than the deterministic model of Guttman scaling shown in Table 3-1 above. The Rasch probabilistic model of the Guttman scaling is the model against which the dataset from item responses of the questionnaire undergoing psychometric testing is evaluated. If the responses show a good fit to the Rasch model the questionnaire that produced them is determined to have met the fundamental principles of measurement for achieving linear (interval level) outcome scores (Tennant and Conaghan, 2007). The relationship between person ability and item difficulty is represented by an S-shaped curve, called an ogive, (Hobart et al., 2007) as shown in Figure 3-2. Note that a change of item difficulty scores from 5 to 10 units and from a score of 45 to 50 units (i.e. a change of 5 units for each) gives a change of person ability on a linear scale that differs by nearly 50% (i.e. from -3 to -1.6 and from -0.2 to 0.5 respectively), illustrating one of the problems with ordinal scores.

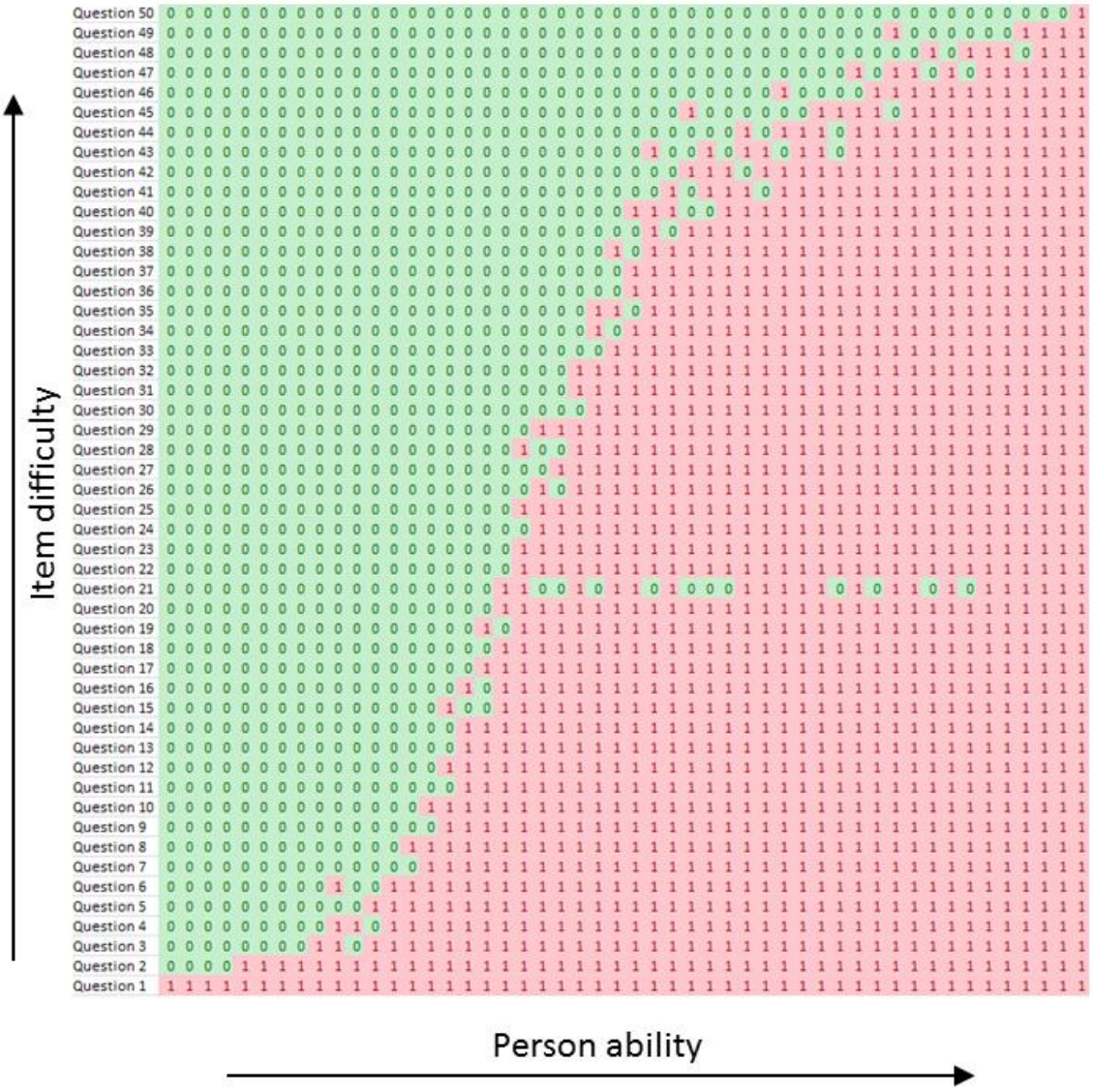
Figure 3-2. The ogive shape representing the relationship between person ability and item difficulty.



The probabilistic Rasch model and the formation of the ogive representing the relationship between person ability and item difficulty are illustrated using a hypothetical educational assessment e.g. a mathematics aptitude assessment. Figure 3-3 below shows the dataset from such a hypothetical assessment, with answers dichotomised into correct answers shown by '1' and incorrect answers shown by '0'. Again, item difficulty is increasing up the y-axis and person ability is increasing along the x-axis. The example assumes that the mathematical aptitude of the sample taking the assessment forms a Normal curve. In this

case, 68% (two standard deviations, one either side of the mean) of respondents' mathematical ability lies equally either side of the mean. This proportion of the sample has no difficulty with passing the least difficult 25 of the 50 mathematical questions, but increasingly struggle with the each successive question. Only the exceptionally gifted mathematicians correctly complete all, or nearly all, of the 50 questions, the final few of which are particularly challenging. Similarly, only mathematically-challenged people at the lowest 5% of ability fail to complete almost all questions. However, for each person, when item difficulty is matched to their ability the probability of passing the question correctly becomes 0.5. This means that there are lucky guesses, carelessness perhaps through haste or confidence, which allow for the occasional 'misfitting' response that presents as a relaxed version of the Guttman pattern. Chi-squared (χ^2) statistics are used to evaluate the fit of the data to the model.

Figure 3-3. The Rasch model's probabilistic version of the Guttman scaling pattern.



Fit to the model indicates that the questionnaire is performing adequately as a measure that meets some of the fundamental requirements that permit the calculation of a linear outcome score from the raw ordinal scores. Failing to endorse the next hardest item allows the person being measured to be numerically quantified on a logistic scale *if the items themselves are on a linear scale, and if they are unidimensional* (they all relate strongly to the characteristic being measured, and not a different underlying characteristic). The linear (interval) scale on which items and persons are numerically located are calibrated in log-odds units called logits; these units represent the natural logarithm of the odds of success i.e. endorsing an item or passing an item (Bond and Fox, 2001). Other psychometric properties must be acceptable for interval-level data scores to be achieved i.e. testing for invariance (stability of items across different levels of ability), Differential Item Functioning

(DIF) and the correct ordering of thresholds of response categories (when items have more than two responses) (Tennant and Conaghan, 2007).

Software used for performing statistical procedures to evaluate fit to the Rasch model (Rasch analysis) allows testing of these properties. It also allows testing for local independence of items through a correlation matrix of residuals – the pattern of differences between the Rasch line of best fit and the actual responses. Local dependence occurs when the response to one item influences the response to another item on the measure, in a similar way to the responses on the hypothetical mobility measure shown in Table 3-1 above. This response dependency was mentioned as a problem with CTT in subsection 3.2 above.

The availability of Rasch analysis software makes Rasch analysis accessible to the wider research community.

χ^2-statistics also show misfitting persons and misfitting items. For example, question 21 shows responses that do not fit with their probability of being successfully passed. This item shows misfit that would be highlighted in the χ^2-statistical analysis, and should be investigated. As an example of why this item does not follow the pattern, see Table 3-2. This shows question 20 and question 21. They are essentially identical questions, but question 21 is presented in a way that challenges interpretation of the question. Perhaps those persons for whom English is not their first language were unable to understand the question; or the question is an example of mathematics which is not the subject of this (hypothetical) mathematical assessment e.g. applied mathematics rather than pure mathematics. In either case, this is an example of the item evaluating other variables, or dimensions, e.g. language, or applied mathematics rather than the branch of mathematics at which the assessment is aimed.

Table 3-2. Questions 20 and 21 from hypothetical mathematics test.

Q. 20	((100 ÷ 20) + (100 ÷ 50)) x 4.50
Q. 21	You hire a van with a full tank of diesel and drive it fully loaded 200 miles. It is quite economical at 40 miles to the gallon. You unload and drive back, at 50 miles to the gallon. How much to fill the van up at £6.20 per gallon.

Items that do not fit the Rasch model affect the fit of the dataset to the model and so must be investigated and examined. In the example given here, it would make sense to remove the item from the questionnaire.

3.2.4 Summary of psychometric methods

Traditional psychometric testing (CTT) has produced the majority of the measures currently used for evaluating upper limb activity limitation. Its strengths are its accessibility due to familiarity of terms, concepts and methods. It has been subject to increasing criticism over recent years because of the untestable nature of, and uncertainty associated with, both its theory and the results of testing measures with CTT methods. Measures developed and tested using CTT produce ordinal level outcome scores which have wide confidence intervals that render them unsuitable for individual use; these scores are unsuitable also for parametric statistical procedures. Potentially, measures developed using CTT have many items; this can produce a misleading high-reliability score and result in local dependence of items which can in turn artificially raise outcome scores. CTT is poor at detecting DIF, and the measure's scores are sample-dependent.

Rasch analysis is less accessible, requiring a greater mathematical understanding of its concepts, but Rasch software is supporting greater use of Rasch methodology across the research community.

Rasch analysis offers a number of advantages over CTT. Firstly, it establishes that when the item responses fit the Rasch model they meet the fundamental requirements of measurement that allow for raw scores to be transformed into interval measurement. Furthermore, Rasch analysis allows evaluation of unidimensionality, DIF and local dependency; more importantly it indicates which items are contributing to the unacceptable psychometrics.

These factors are important considerations when evaluating the acceptability of measures for use in research and clinical practice.

3.3 Critical review of measures of upper limb functional ability for children with cerebral palsy

3.3.1 Introduction

Outcome measures used to evaluate changes in activity limitation in children with cerebral palsy engaging with experimental approaches such as CIMT and botulinum toxin are not always appropriately validated, reliable or responsive for the purpose in which they are intended for use (Hoare et al., 2007a, Hoare et al., 2010, Hobart et al., 2007, Tennant, 2007). It is difficult to accept at face value the results of trials or any rehabilitation programmes even if their conduct is diligently performed and scientifically sound when the measures generating the results are psychometrically poor. As recently as June 2014, Geerdink et al. (2014) suggested that there are no validated measures for evaluating upper limb activity limitation of the affected upper limb in bimanual activities.

Developments in psychometric techniques used for evaluating scales (e.g. Rasch analysis) have raised questions about the psychometric properties of measures developed using CTT (Hobart et al., 2007). For example, a measure's reliability can be improved by increasing the number of items. However, this can result in local dependency of items, masking a loss of unidimensionality and falsely inflating scores to give potentially misleading results (Tennant and Conaghan, 2007). Furthermore, arithmetic operations on ordinal level data are inappropriate, making it difficult to evaluate outcomes and demonstrate efficacy of treatment programmes (Grimby et al., 2012, Hobart et al., 2007, Linacre, 2000, Stucki et al., 1996, Tennant and Conaghan, 2007). The wide confidence intervals and non-linear nature associated with ordinal-level data are persuasive reasons given by Hobart et al. (2007) for using only outcome measures which produce interval level data. Finally, if a measure is to meet the fundamental principles of measurement, evidence of acceptable unidimensionality, absence of DIF and a stability of item difficulty across varying person ability are essential and must be demonstrable (Tennant and Conaghan, 2007).

The development of any measure involves careful consideration of the trait and the population being measured (Wilson, 2005). This means that measures should be

developed specifically for the patient population and the health condition under evaluation. Generic measures (those that claim to measure a specific trait across a broad range of health conditions) can be unresponsive to changes in the trait being measured (Arnould et al., 2012, McCullough and Parkes, 2008). Norm-referenced measures are only suitable as discriminative tools (Palisano et al., 1995, Rosenbaum et al., 1990) unless they have been validated as responsive in the population for which they being used (Rosenbaum et al., 1990, Centre for Reviews and Dissemination, 2013).

In order to maximise the robustness and quality of the findings of this research project an independent and meticulous investigation was performed to identify a psychometrically sound outcome measure appropriate for evaluating changes in upper limb activity limitation of children with cerebral palsy. Firstly, this involved a systematic search following procedures described by the Centre for Reviews and Dissemination (Centre for Reviews and Dissemination, 2013). The aim of this systematic search was to identify from the literature every outcome measure that is being used to evaluate upper limb function of children with cerebral palsy. Details were then collated from the search results of each measure's type of outcome score data, reliability, validity, dimensionality and responsiveness. Evidence of its clinical utility – ease of scoring, ease of use, training and costs – were also noted.

This was not a systematic review, therefore, but a critical appraisal (Grant and Booth, 2009) of every measure used in research (and potentially by clinicians) as a means for evaluating upper limb activity limitation (functional outcomes). This appraisal did not reject any papers if they were poor quality, but used these papers to identify whether the measure in question was poorly tested and developed. Only one researcher identified the relevant papers and evaluated the quality of evidence for the measures identified in the search. Although this does not reach the scientific quality of a systematic review, it is a recognised methodology and review for the purpose (Grant and Booth, 2009) of locating relevant papers in order to identify the psychometric weaknesses and strengths of all measures used for evaluating changes to upper limb activity limitation in children with cerebral palsy. For this reason, measures that evaluate unilateral hand function are included alongside those that evaluate bilateral hand function. The aim of this appraisal is to identify measures that have achieved acceptable psychometric standards for the

evaluation of upper limb activity limitation of children with cerebral palsy and to describe those outcome measures' characteristics, strengths and weaknesses.

3.3.2 Methodology

The search for measures of activity limitation for children with cerebral palsy was performed on the established databases for this purpose, AMED, EMBASE, Ovid MEDLINE, using the search strategy described by the Centre for Reviews and Dissemination (Centre for Reviews and Dissemination, 2013). The search strategy is shown in Table 3-3 and was last performed in August 2013.

Table 3-3. Search strategy for critical review of measures of activity limitation.

Databases searched: AMED, EMBASE, Ovid MEDLINE
1. cerebral palsy
2. CP
3. 1 OR 2
4. child*
5. p?ediat*
6. 4 OR 5
7. Functional Independence Measure
8. WeeFIM
9. Jebsen-Taylor
10. Jebsen
11. P?ediatric Evaluation Disability Inventory
12. PEDI
13. Assisting Hand Assessment
14. Jebsen Hand function test
15. Bruininks-Oseretsky
16. ABILHAND-kids
17. Motor Activity Log
18. Paediatric Motor Activity Log
19. Pediatric Motor Activity Log
20. PMAL
21. Peabody Developmental Motor Scales
22. PDMS
23. Melbourne Assessment Unilateral Upper Limb Function
24. Quality Upper Extremit* Skills Test
25. QUEST
26. BOTMP
27. Canadian Occupational Performance Measure
28. COPM
29. GAS
30. Goal Attainment Scal*
31. measur*
32. evaluat*
33. assess*
34. scale
35. 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34
36. reliab*
37. respons*
38. sensitiv*
39. vali*
40. construct
41. psychometric*
42. 36 OR 37 OR 38 OR 39 OR OR 40 OR 41
43. function*

44. activit*
45. Activit*
46. 43 OR 44 OR 45
47. upper limb
48. arm
49. upper extrem*
50. hand
51. 47 OR 48 OR 49 OR 50
52. 3 AND 6 AND 35 AND 42 AND 46 AND 51

Medline searches for this word plus any combination of characters e.g. reliab produces a search for reliability and reliable
? signifies a single character only or the possibility that no character exists there e.g. p?ediatric produces a search for paediatric and pediatric.

These search terms included outcome measures that were familiar to the research team through team members' clinical practice and through their use in studies which were part of the team's evidence base when preparing this research study. The search elicited a total of 5,470 articles from the three databases. Removing duplicates left 4,375 articles. The remaining titles and abstracts were examined for content. All articles were retained that reviewed or evaluated the characteristics and psychometric properties of measures of upper limb activity limitation for children with cerebral palsy, or which involved evaluating upper limb activity limitation of children with cerebral palsy e.g. Wallen et al. (2007). The only criteria for exclusion therefore were articles which had no information relevant to the aims of this critical appraisal. Retained articles were reviewed for the methodology used to establish the psychometric properties of the measures, or for their responsiveness and descriptions of their use. A secondary search was performed to identify any measures that were mentioned in articles identified in the primary search if those measures' properties were not fully described. Where characteristics (e.g. assessment time, scoring etc.) could not be ascertained from the published article, information was obtained from instruction manuals or websites of the measures. This process produced 141 articles and identified 21 measures, from which information about the measures' characteristics, use and psychometric properties was extracted. Table 3-4 below lists the measures, their characteristics and other background information.

Table 3-4. The 21 measures located by the search strategy, their characteristics and other background information.

Measure	Population validity	Domains	Time taken (mins) to administer and score	Number of items	Summary of findings and further information	Cost
Canadian Occupational Performance Measure (COPM)	All ages, validated in many populations including children with cerebral palsy.	Non-standardised measure, generates individual goals across the ICF-CY Activity and Participation domains in a parent-therapist discussion.	20-40 minutes (Sršen, 2012), through semi-structured interview to develop up to five non-standardised goals. Parent rates each goal out of ten for perceived performance and satisfaction for performance.	Ordinal outcome scores (derived using questionable mathematical procedures) (Tennant, 2007).	The COPM is an apparently responsive measure but its psychometric properties are poor because of the use of ordinal data within its mathematical scoring. The lack of unidimensionality and the ordinal outcome scores mean that outcome scores should not be used with non-parametric analyses and should be interpreted with caution (Stucki et al., 1996, Tennant, 2007).	From CAN$52.45 for 100 forms.
ABILHAND-kids	Parent-reported questionnaire validated for children with cerebral palsy.	Activity and Participation.	Administration: 5 – 10 minutes. Scoring: 10 minutes.	Twenty-one manual activities scored using a three-level scale (impossible, difficult, easy). Online scoring for conversion to interval scale data.	Similar in principal to the ABILHAND (for adult stroke patients) the ABILHAND is Rasch-derived with excellent psychometric qualities. It presents summed and interval data outcome scores, and is reported as one of the best psychometric and clinical	Freely available for download.

					measures for its purpose (Gilmore et al., 2010) . There are doubts about its responsiveness, and it has floor and ceiling effects.	
Bruininks-Oseretsky Test of Motor Proficiency (BOTMP) http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=PAa58000	Children and adults with motor impairment aged four to twenty-one years old. Unsuitable for children with cerebral palsy and for pre-school children. Norm-referenced.	Eight domains, four relevant to upper limb: Fine Motor Precision. Fine Motor Integration. Manual Dexterity. Bilateral Coordination.	40 – 80 minutes to administer (Deitz et al., 2007). Scoring takes ten minutes but is challenging, with errors likely (Deitz et al., 2007).	Upper limb: Up to 27 items in four subtests (but bilateral tests include lower limb testing).	Scoring and interpretation of scores is challenging (Deitz et al., 2007), the BOTMP is norm-referenced and there is no evidence for its use in children with cerebral palsy.	$799.00
Goal Attainment Scale (GAS)	Disabled children including cerebral palsy.	Non-standardised measure, generates individual goals across the ICF-CY Activity and Participation domains in a parent-therapist discussion.	Reported as time-consuming (Steenbeek et al., 2007) .	Up to five goals rated on five point ordinal scale, using questionable mathematical procedures Questionable scoring practices converts to mean normalized T score of 50.	The GAS is an apparently responsive measure but its psychometric properties are poor because of the use of ordinal data within its mathematical scoring. The lack of unidimensionality and the ordinal outcome scores mean that outcome scores should not be used with non-parametric analyses and should be interpreted with caution (Stucki	Free

					et al., 1996, Tennant, 2007).	
Paediatric Motor Activity Log (PMAL)	Tested using Rasch analysis and validated for children with cerebral palsy (but requires further psychometric testing).	Two scales: 'How Often' and 'How Well'.	Administration : 5 - 15 minutes (Wallen et al., 2009b) Scoring: 3 minutes (Wallen et al., 2009b).	Twenty-two bilateral and unilateral functional items, scored on a three point ordinal scale (following Rasch-derived modification from five point scale (Wallen et al., 2009b).	A potentially useful measure, but requires further psychometric testing of the response categories. Responsiveness of the Quality of Use is good.	Free
Paediatric Evaluation of Disability Inventory (PEDI)	Generic, chronically ill and disabled, not specifically for children with cerebral palsy but has been validated for this population (McCarthy et al., 2002, Nichols and Case-Smith, 1996).	Two dimensions: functional skills scales; carer assistance scales. Three domains: self-care, mobility, social function.	45-60 minutes through Direct observation or structured parent interview(Klingels et al., 2010).	197 functional skill items. Items graded Unable or Capable for functional capability. (limited upper limb assessment of <35%).	Inconclusive or poor responsiveness, but has had substantial validation studies performed. The Rasch analysis was carried out on data from non-disabled children, The PEDI is norm-referenced.	$124.10 manual, $143.60 for 100 forms.
Peabody Developmental	Generic, norm-referenced	Six subtests: Reflexes, Stationary	20-30 minutes (Klingels et al.,	Fine-Motor Scale contains 112 items	Norm-referenced with poor responsiveness and no evidence	$455.00 25 copies.

Motor Scales (PDMS)	(Palisano et al., 1995) .	(body control and equilibrium), Locomotion, Object Manipulation, Grasping, and Visual-Motor Integration.	2010).	evaluated on a three-point scale.	for its validity as an outcome measure in children with cerebral palsy.	
Melbourne Assessment of Unilateral Upper Limb Function (MAUULF)	Children with cerebral palsy, criterion referenced.	Scores quality of unilateral UL function: reach, grasp, release, manipulation)(most items unrelated to Activity, the majority related to Body Functions (Hoare et al., 2011). range, accuracy, fluency and dexterity.	30 mins (Gilmore et al., 2010) Video recorded. Scoring – 30 minutes (Klingels et al., 2010). Scores significantly higher in trained users (Cusick et al., 2005).	Sixteen items scored on a three, four or five point scale. Raw scores to percentages.	Reported as having the best psychometric evidence for use with unilaterally impaired children with cerebral palsy (Gilmore et al., 2010), there are questions about its responsiveness. Less than half of its items relate to upper limb Activity, and Sakzewski et al. (2011b) suggest this is related to its poor responsiveness. Validation has only been performed for Australian children. A Rasch analysis has identified good properties generally but there are nine of 35 items (25.7%) displaying poor reliability that needs investigating (Randall et al., 2010).	$927.00
Quality of Upper	Children with	Dissociated	15 minutes	34 items: (Hoare et	Only a third of items relate to	Freely available for

Extremity Skills Test (QUEST)	cerebral palsy aged 18mo– 8y, criterion-referenced.	movements, grasp, weight bearing, protective extension. Only Grasp domain measures Activity (Hoare et al., 2011).	(Klingels et al., 2010). Scoring 15 – 30 minutes (Klingels et al., 2010).	al., 2011) Dissociated movement has 19 items, only four related to Activity Grasp: 15 items, 14 related to Activity Weight-bearing: 26, only two related to Activity Protective Extension: 18 of which none were Activity. Between two and six responses for each item, raw scores converted to percentages. Interpretation of current scoring system misleading, 14 items are a poor fit to the Rasch model (Thorley et al., 2012a).	Activity, and there are questions about its responsiveness. Outcome score is ordinal, despite the discouraged practice involved with its averaged ordinal score conversion to a percentage (Tennant, 2007). Substantial further psychometric testing and reorganisation of scoring is necessary, with possibility of removal of items and even domains (Thorley et al., 2012a).	download

Assisting Hand Assessment (AHA)	Children who have a unilateral disability, including children with hemiplegic cerebral palsy. Rasch derived.	Assisting hand general use, arm use, grasp-release, fine motor adjustment, coordination and pace (Gilmore et al., 2010).	15 minutes play-session (10 – 15 minutes (Krumlinde-Sundholm et al., 2007)), video recorded and scored afterwards (takes from15–30 minutes (Klingels et al., 2010)) up to one hour (Gordon, 2007).	22 test items, four-point criterion-referenced scale from 22 points, meaning that the hand is not used at all, to 88 points. Scoring and interpretation are complex (Krumlinde-Sundholm et al., 2007) .	A well-developed measure, but requires extensive training to use. Developed using Rasch analysis, two reviews find this measure one of the most psychometrically sound measures for the evaluation of bilateral upper limb Activity in children with unilateral cerebral palsy (Gilmore et al., 2010, Greaves et al., 2010).	Expensive in terms of money and time: participants undertake three day training course and must also complete eight calibration cases.
Assessment of Motor and Process Skills (AMPS) Generic measure, norm-referenced, has had substantial psychometric testing including Rasch analysis , but designed primarily to	Generic, norm referenced, designed to measure quality of a person's Activity performance (Fisher and Merritt, 2012). Validity sample =	No domains, but extensive list of 120 functional activities from which two are selected and performed in usual environment.	30-40 minutes (http://strokengine.ca/assess/module_amps_quick-en.html)(http://www.innovativeotsolutions.com/content/amps/be	One to two tasks (Activities) selected from one of 120 in the AMPS list. Rater scores quality of performance on each of 16 motor and 20 process items (i.e., occupational performance skills)	Expensive, with extensive training required, a generic norm-referenced measure which measures self-care rather than upper limb Activity.	£800, five day training and calibration course, with follow-up testing of 10 people. Scoring software costs $199.

measure self-care, not upper limb Activity (Gilmore et al., 2010).	2 024 children with neurological developmental disorders including spina bifida, multiple unspecified developmental disorders (e.g. cerebral palsy) (Fisher and Merritt, 2012).		nefits/)	according to the standardized criteria in the AMPS manual. Each task performance observed is scored separately and each ADL skill is rated using a four-point ordinal scale. AMPS software used to generate linear outcome score from ordinal raw items scores on quality of task performance.		
Jebsen–Taylor Test of Hand Function (JTT)	No psychometric testing on children with cerebral palsy known (Gilmore et al., 2010) . Not indicated as reliable or valid in children with cerebral palsy (Lemmens et al., 2012).	Evaluates a range of unimanual hand activity. http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm.aspx?ID=1025	5 – 30 minutes, scored by summing times for each subtest. Items attempted with each hand using non-dominant hand first. http://www.rehabmeasures.	Seven subtests (performed on both non-dominant and dominant hand) : Writing a 24-letter sentence (eight year old reading difficulty) Card turning Picking up small common objects and placing them in a container Stacking checkers	Measures unilateral hand function through speed, not quality of performance. There is no evidence for psychometric testing, reliability or validity for testing children with cerebral palsy using the JTT.	$428.05 from Amazon.com.

	Norm referenced (Klingels et al., 2010).		org/Lists/RehabMeasures/DispForm.aspx?ID=1025	Simulated feeding Moving light objects (empty cans) Moving heavy objects (weighted cans (1 lb)) . http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm.aspx?ID=1025		
The Functional Independence Measure for Children (WeeFIM)	Generic measure for chronically ill children with disabilities (Debuse and Brace, 2011).	Six domains: self-care, sphincter control, transfers, locomotion, communication, social cognition (Ottenbacher et al., 1996).		18 items on a seven-level ordinal scale. Scores are summed.	Not valid for use with children with cerebral palsy due to high ceiling effects and poor responsiveness (Debuse and Brace, 2011). Designed primarily to measure self-care, not upper limb Activity (Gilmore et al., 2010).	
Shriners Hospital Upper Extremity Evaluation (SHUEE)	Validated only for American children with unilateral involvement (Gilmore et al.,	40 items (Lemmens et al., 2012) in two sections: first examines several aspects of upper limb performance	15 minutes, therapist administered in a standardised format, scored	Scores based on percentage of maximum score for each section.	There is no evidence for tests of unidimensionality, and scoring and variety of analyses suggest that the outcome score includes aspects of upper limb performance, Body Structures	Free online

	2010).	e.g. range of movement, spasticity, ADL performance ; second section has three parts involving analysis of spontaneous function, dynamic positioning and grasp and release (Gilmore et al., 2010).	afterwards (Gilmore et al., 2010). Scoring takes 15 – 30 minutes (Klingels et al., 2010).		and Body Functions. The validity and reliability testing has been carried out on children who have had orthopaedic surgery (Davids et al., 2006) and using an unacceptable validation measure (Gilmore et al., 2010).	
House Functional Classification (HFC) including the modified HFC (mHFC)	Validated for upper limb evaluation children with cerebral palsy undergoing surgery, specifically for Thumb in Palm deformity.	Nine grades of hand use ranging from 0 (Does not use) to 8 (Spontaneous use, complete) (Koman et al., 2008). Each grade has a number of descriptors totalling 32 (Koman et al., 2008).	Undefined.	Scores appear based on summation of the descriptors (Koman et al., 2008).	No modern psychometric testing, no tests of unidimensionality, evidence of poor responsiveness.	Free with access to paper by Koman et al (Koman et al., 2008).
Pediatric Outcomes Data Collection Instrument (PODCI)	A generic instrument with limited validation studies carried	Four domains: Upper Extremity Function, Transfers and Basic Mobility.	Undefined.	114 items altogether (Barnes et al., 2008), scored between 0 – 100 using a	Psychometric evidence and findings of reviews conclude that the PODCI is not a valid measure of function for CP population.	Free for download from the website of The American Academy of Orthopaedic Surgeons

	out for upper limb activity (Barnes et al., 2008) including a Rasch analysis (Seok et al., 2012) because many of the children potentially had no upper limb problems.	Sports and Physical Function,, and Comfort/Pain (Barnes et al., 2008). Global function is an average of the four scores (Lerman et al., 2005).		Microsoft Excel spread sheet freely available from the website of The American Academy of Orthopaedic Surgeons http://www.aaos.org/research/outcomes/outcomes_peds.asp		http://www.aaos.org/research/outcomes/outcomes_peds.asp
Nine hole peg test (NHPT) and the Box and Block test (BBT)	The only validation studies carried out in children with cerebral palsy aged 7.6 years to 13.4 years old (Van Hedel and Wick, 2011).	N/A	Less than five minutes, but depends on level of disability.	N/A	No evidence of construct validity, reliability or relevance to Activity Limitation in children with cerebral palsy. .	The NHPT is free apart from purchase of materials (if self-constructed, see http://www.rehabmeasures.org/lists/rehabmeasures/dispform.aspx?id=925) The BBT costs $200.00 (http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm.aspx?ID=917)
Video Observations Aarts and Aarts	Children with unilateral	N/A	20 to 40 minutes (Aarts	Two standardised tasks (stringing	Requires attendance at a three-hour training course (Aarts et al.,	Costs €500 including training course and

(VOAA)	cerebral palsy aged 2.5–10 years (Aarts et al., 2009, Aarts et al., 2007).		et al., 2009), but needs computer software to assist (Aarts et al., 2009). Tasks take 2 – 7 minutes, average scoring took 30 minutes (Houwink et al., 2013).	beads, decorating a muffin), each with four subtasks (Houwink et al., 2013) which can be used to evaluate reaching, grasping and holding (Aarts et al., 2007).	2009). Requires video recording and use of software to analyse and score.	software http://www.freewebs.com/voaa/education.htm
Assessment of Children's Hand Skills (ACHS) related to the CHSQ (see below)	A generic instrument, validated only in Taiwanese children and evaluated using Rasch analysis but needs further psychometric testing (Chien et al., 2011b).	Three domains (leisure and play, school/education, and activities of daily living) for 22 Activities, these correspond to the Activities in the parent-reported companion measure, the CHSQ (Chien et al., 2011b). The assessment is based on	Each of two to three activities are observed for a maximum of ten minutes. A six-level rating scale is used to rate the hand skills in each Activity. (Chien et al., 2011b).	Six-level rating scale on two or three of 22 activities that are selected after parent-reported assessment using the Children's Hand-Skills ability Questionnaire (CHSQ) (Chien et al., 2011b).	No psychometric testing performed for validation in the UK and validation studies to be performed in the UK. Needs further psychometric testing and development (Chien et al., 2011b).	Free

		observation of 20 hand skill items in six domains: manual gesture, body contact hand skills, arm-hand use, adaptive skilled hand use, bilateral use, and general quality (Chien et al., 2011b).				
Children's Hand-Skills ability Questionnaire (CHSQ) related to the ACHS as a companion assessment.	A generic measure, validated only in Australia and Taiwanese children using Rasch analysis (Chien and Brown, 2012).	22 Activities in three domains: leisure and play, school/education, and activities of daily living (Chien and Brown, 2012) corresponding to the ACHS.	Not reported.	Three standard responses for parents to report level of difficulty for the 20 hand skills in each of 22 activities (Chien and Brown, 2012).	Needs further development (Chien and Brown, 2012). No psychometric testing performed for validation in the UK and validation studies to be performed in the UK.	Free
Children's Hand-use Experience Questionnaire (CHEQ)	Validated for children with unilateral impairment:	Includes three scales: perceived efficacy	Time taken is not described. Completion is performed	After Rasch analysis, 29 items (activities) are included in this questionnaire (Skold	This measure has a solid basis for its development, including Rasch analysis to evaluate its psychometric properties, but it	Free to use and obtain a summary of the ratings, but a score requires software for which there

	brachial plexus injury, cerebral palsy and upper limb reduction deficiency (Skold et al., 2011).	of the grasp; time taken to perform the activity; degree of feeling bothered.	online (http://www.cheq.se/) and a free report summarising the ratings is available. There is a charge for the software which is necessary to score the questionnaire.	et al., 2011). Each activity has five questions, two with three responses and three with four responses. These questions and responses are standard for each activity.	requires further testing after its modifications. Validated only in Sweden.	is an unspecified fee.

COPM : Canadian Occupational Performance Measure; ICF-CY: International Classification on Function, Disability and Health for Children and Youth; BOTMP: Bruininks-Oseretsky Test of Motor Proficiency; GAS: Goal Attainment Scale; PMAL: Paediatric Motor Activity Log; PEDI: Paediatric Evaluation of Disability Inventory; PDMS; Peabody Developmental Motor Scales ; MAUULF: Melbourne Assessment of Unilateral Upper Limb Function; QUEST: Quality of Upper Extremity Skills Test; AHA: Assisting Hand Assessment; AMPS: Assessment of Motor and Process Skills; JTT: Jebsen–Taylor Test of Hand Function; WeeFIM: The Functional Independence Measure for Children; SHUEE: Shriners Hospital Upper Extremity Evaluation; ADL: Activities of Daily Living: HFC: House Functional Classification; mHFC: modified House Functional Classification; PODCI: Pediatric Outcomes Data Collection Instrument; NHPT: Nine hole peg test; BBT: Box and Block test; VOAA: Video Observations Aarts and Aarts; ACHS: Assessment of Children's Hand Skills; CHSQ: Children's Hand-Skills ability Questionnaire; CHEQ: Children's Hand-use Experience Questionnaire.

Many of the articles, such as the systematic reviews (Sakzewski et al., 2007, Harvey et al., 2008, Gilmore et al., 2010, Greaves et al., 2010, Klingels et al., 2010, Debuse and Brace, 2011, Lemmens et al., 2012) contained information about several of the 21 measures and were particularly relevant and helpful to the critical appraisal. All data regarding each measure were extracted and collated. Some measures are designed and validated for bilateral use; others focus only on the impaired upper limb of children with unilateral impairment. Because the aim of this paper is to appraise the literature relating to all measures of upper limb activity limitation, papers relating to both types of measure were included in the appraisal.

Articles that used modern techniques such as Rasch analysis to test the measures' psychometric properties were not favoured over articles using CTT; each measure's psychometric properties were extracted and collated for the purposes of comparison and openness in the appraisal process. In Rasch analysis, reliability is usually shown as the Person Separation Index (PSI) and in CTT as Cronbach's alpha and both are given where known. Because the number of items inflates reliability (Tennant and Conaghan, 2007), the number of items in each measure are given.

3.3.3 Results

Once the properties and characteristics of the measures were collated, the data were reviewed by the Chief Investigator (NP) and a senior academic supervisor. The psychometric properties of the 21 measures and the assessment of their strengths and weaknesses are presented in Table 3-5 below.

Table 3-5. Psychometric properties and details of the 21 measures.

	Targeting			**Reliability**		**Validity**	**Responsiveness**	
Name of measure	Age and population validation	Floor effect	Ceiling effect	Test-retest	Inter-rater		Responsiveness Indicated by SRM or effect size where available) (Lindeboom et al., 2005)	MCID
Canadian Occupational Performance Measure (COPM) A clinically-appropriate and apparently responsive measure, but the scoring system is mathematically and psychometrically unsound. This has the potential to increase the issues with ordinal outcome scores (Tennant, 2007, Grimby et al., 2012)	All ages. Validated for many conditions including cerebral palsy and paediatric use.	Not reported.	Not reported.	Spearman's rho correlation coefficient 0.76–0.89 (Sakzewski et al., 2007) but these values are from studies on adult patients with stroke and chronic lung disease.	Described as moderate to good, but no figures given (Verkerk et al., 2006).	Internal consistency alpha = 0.73 for performance (Cusick et al., 2007).	Better than stated MCID in manual (Cusick et al., 2007) Good responsiveness supported in five studies (Sakzewski et al., 2007).	Change of 2.0 on the subjective score rated out of 10 (COPM manual).

ABILHAND-kids Parent-reported criterion referenced questionnaire, psychometrically tested using Rasch analysis and able to produce interval data outcome scores. Validated on a Belgian population using a French version before translation. Requires validation studies in other countries.	6 - 15 years old. Validated for use with children with cerebral palsy in Belgium (Arnould et al., 2004).	For the most severely affected children i.e. MACS V. Tested on only a few children with severe disability (MACS IV)(Arnould et al., 2004).	No (Arnould et al., 2004).	Pearson correlation coefficient = 0.91 (Arnould et al., 2004).	Not reported.	High (Rasch-derived 0.98) (Arnould et al., 2004).	Not reported (Lemmens et al., 2012). Poor responsiveness demonstrated in adult version (Bovolenta, 2009).	Not known. A difference of the standard error of the logit scores (approximately 0.45) was used to indicate a clinically significant change
Bruininks-Oseretsky Test of Motor Proficiency (BOTMP) Norm-referenced, no evidence of validation in children with cerebral palsy.	4 – 21 years old No evidence of validation for its use in children with cerebral palsy.	Yes (Venetsanou, 2007).	Not reported.	Pearson correlation coefficient = 0.86 (subtest eight, Fine Motor skills) (Klingels et al., 2010) Pearson product moment correlation coefficients	Pearson correlation coefficient = 0.94 (subtest eight, Fine Motor skills) (Klingels et al., 2010). Pearson product moment correlation coefficient = 0.86 for the Fine Motor	Four types described: Content validity, internal structure, differentiation between clinical and non-clinical groups, convergent validity (Deitz et al., 2007). Problems with construct validity	Poor (Venetsanou, 2007).	Not reported.

				<0.8 for subtests (Deitz et al., 2007). Inconsistent and some limitations indicated (Wilson, 2000)	Precision subtest (Deitz et al., 2007).	indicated (MacCobb et al., 2005) .		
Goal Attainment Scale (GAS) A clinically-appropriate and apparently responsive measure, but the scoring system is mathematically and psychometrically inappropriate, potentially leading to outcome scores which are very inaccurate (Tennant, 2007).	All ages and conditions.	Yes but avoidable (Steenbeek et al., 2007).	Not reported.	Not reported.	kappa coefficient = 0.89 (Steenbeek et al., 2007).	Ambiguous (Steenbeek et al., 2007).	Good responsiveness (Steenbeek et al., 2007) .	Not reported but the scoring system suggests that any change in outcome scores would mean a clinically significant change.
Paediatric Motor Activity Log (PMAL) Criterion-referenced measure, psychometrically tested using Rasch analysis but which needs further	Unilateral CP aged six months – eight years in Australia (Wallen et	Floor and ceiling effects for extremes of ability, possibly marked		ICC = 0.94 and 0.93 (How Often and How Well, respectively	Not reported.	Person separation index 0.9 (Wallen et al., 2009b). Concurrent	Good responsiveness observed in one study (DeLuca et al., 2006) Good responsiveness SRM = 0.89–0.99 (Lin	0.5 SD of the baseline score (Lin et al., 2012): a change of 0.67 points or

testing after reduction of item response categories (Wallen et al., 2009b). Lin et al (2012) performed validity and responsiveness testing but used inappropriate measures for testing validity	al., 2009b) and further validation study preformed in Taiwan (Lin et al., 2012).	effects (Wallen et al., 2009b). Easiest item likely to be inappropriate for children aged six and older, increasing the floor effect.	) (Wallen et al., 2009b).	validity tested against the weeFim and the PMAL (Lin et al., 2012), both of which this paper rules out for evaluating activity of children with cerebral palsy. The Amount of Use scale showed poor validity (Lin et al., 2012).	et al., 2012).	higher on Amount of Use, and 0.66 points on Quality of Use.

Paediatric Evaluation of Disability Inventory (PEDI) Norm referenced using non-disabled children and developed for children with disabilities (Nichols and Case-Smith, 1996), has had substantial psychometric testing including Rasch analysis (but Rasch analysis on non-disabled children's data (McCarthy et al., 2002)).	Six months – seven years old. Generic instrument for variety of disabilities, not reliability or validity tested for children with cerebral palsy (McCarthy et al., 2002) but psychometric testing performed on a very small sample of children with cerebral palsy (Nichols	None (McCarthy et al., 2002).	Not reported (McCarthy et al., 2002) Ceiling effect reported in more able children (Vos-Vromans et al., 2005).	Test-retest ICC = 0.91–0.98 (Iyer et al., 2003).	ICC = 0.99 (Iyer et al., 2003).	Concurrent, discriminant and evaluative validity has been carried out but not on children with cerebral palsy (McCarthy et al., 2002). Rasch-developed but using data from healthy children (McCarthy et al., 2002). Validity testing only a very small study.	Inconclusive (variable reports from studies (Harvey et al., 2008) . Unresponsive in clinical trial (Russo et al., 2007).	11 points on the transformed PEDI score (11%) (Iyer et al., 2003).

	and Case-Smith, 1996).					
				internal consistency α = 0.98 (McCarthy et al., 2002).		
Peabody Developmental Motor Scales (PDMS) Requires significant revision if it is to meet adequate psychometric standards in a measure on Taiwanese children (Chien and Bond, 2009) .	Birth to seven years, based on children following healthy developmental course.	Not reported.	Significant ceiling effect in Taiwanese children (Chien and Bond, 2009).	Lack of adequate data on validity (Gilmore et al., 2010) .	Poor responsiveness (Palisano et al., 1995). Some testing to improve its use as an evaluative measure in children with CP has been carried out in Taiwan, on very young	Not reported.

Norm-referenced, should not be used as an evaluative measure (Palisano et al., 1995).							children (1 – 4.5 years) and is sensitive to change with intervention over three months (Wang et al., 2006). Not responsive in clinical trial(Russo et al., 2007).	
Melbourne Assessment of Unilateral Upper Limb Function (MAUULF) Aim of MAUULF is to measure the quality of upper limb Activity (Hoare et al., 2011, Johnson et al., 1994), validated only for use of the affected arm (Gilmore et al., 2010). Scored from a video, criterion-referenced, less than half the items relate to Activity (Hoare et al., 2011, Gilmore et al., 2010). A Rasch analysis has confirmed some psychometric strengths, including unidimensionality	Five to fifteen years (Klingels et al., 2010), validated only in an Australian population. Further validation study supports validity from six months in an Australian population (Randall et	Not reported.	Not reported.	Moderate to high (ICC = 0.79 and 0.83) (Randall et al., 2001).	ICC = 0.95 (Randall et al., 2001)and 0.98 / 0.99 in untrained / trained groups respectively (Cusick et al., 2005).	Cronbach's *α* = 0.96 (Randall et al., 2001) and 0.99 (Cusick et al., 2005). Concurrent validity = 0.939 (with PEDI self-care) (Bourke-Taylor, 2003). Excellent Person Separation Index ≥ 0.92 (Randall et al., 2010).	Non-responsive in clinical trials: (Wallen et al., 2007), in which the GAS and COPM showed functional improvements; Speth et al (2005); Sakzewski et al. (2011b) also suggest non-responsiveness . Scores significantly higher in trained users (Cusick et al., 2005).	9% (Klingels et al., 2008).

of each scale, but there are problems with nine items (of 35, 25.7%) which need addressing (Randall et al., 2010), and is validated only in an Australian population.	al., 2008).							
Quality of Upper Extremity Skills Test (QUEST) Criterion-referenced, only 35% of items contribute to activity (Gilmore et al., 2010), validated only on an Australian population.	Eighteen months to eight years (Klingels et al., 2008) and only in an Australian population.	Not reported.	Not reported.	ICC = 0.95 (DeMatteo et al., 1993) . ICC = 0.945 (Thorley et al., 2012b) . Norwegian study shows poorer reliability in the Grasp domain (only domain which reflects Activity) ICC = 0.68 – 0.78(Sorsda	ICC = 0.75 – 0.95 (DeMatteo et al., 1993), validated only on an Australian population. ICC = 0.861 (Thorley et al., 2012b). Norwegian study shows poorer reliability in the Grasp domain (only domain which reflects	Concurrent validity, Pearson's product moment correlation coefficient = 0.84 (with PDMS, total scores)(DeMatteo et al., 1993). Internal consistency (*a* = 0.976) (Thorley et al., 2012b) but authors recognise that number of items positively influence alpha.	Non-responsive in a clinical trial (Wallen et al., 2007)in which the GAS and COPM showed functional improvements.	13.8% (for the affected upper limb)(Klingels et al., 2008).

				hl et al., 2008).	Activity) ICC = 0.69(Sorsdahl et al., 2008). A recent Rasch analysis (Thorley et al., 2012a) showed poor psychometric properties (unidimensionality only within domains, dependency of items, recommendations with scoring to prevent misleading interpretation of results).			
Assisting Hand Assessment Rasch-developed, criterion-referenced measure, validated for Swedish children with cerebral palsy who have unilateral impairment.	Swedish children aged eighteen months to twelve years with unilateral impairment (Krumlinde-Sundholm et al., 2007).	Not detected (Krumlinde-Sundholm et al., 2007).	Not detected (Krumlinde-Sundholm et al., 2007).	ICC = 0.98 - 0.99 (Gilmore et al., 2010).	ICC = 0.97 – 0.98 (20 raters (Gilmore et al., 2010)).	Person separation index (Rasch-derived equivalent of Cronbach's) = 0.97 (Krumlinde-Sundholm et al., 2007) . Validated for a Swedish population.	Evidence of responsiveness in a clinical trial (Eliasson et al., 2005) The high person separation index of 0.97 indicates good responsiveness (Krumlinde-Sundholm et al., 2007).	Four sum score point (Holmefur et al., 2009) .
Assessment of Motor and Process Skills (AMPS) Generic norm-referenced	Three to fifteen years, and	No floor or ceiling effects evaluated or		Detailed description of various reliability and validity testing on numerous populations (Fisher and Merritt, 2012).			Not responsive in clinical trial with children with	Not given: trials have looked for an

measure of quality of performance (Fisher and Merritt, 2012), tested post-development using Rasch analysis (Boyd et al., 2012) . Designed primarily to measure self-care, not upper limb Activity (Gilmore et al., 2010).	ten age groups between sixteen and 103 years. Norm referenced, generic measure.	reported, awareness of these problems noted by developers, and development of AMPS designed to minimise them (Fisher and Merritt, 2012).		Rasch-derived reliability (high inter- and intrarater reliability: mean square value ≤ 1.4, $z < 2$) using 13,070 occupational therapists. Test-retest reliability coefficient = .87 (on elderly patients). Person Separation Index (Rasch equivalent of Cronbach's alpha) = 0.92. Validity: 90.5% of 15,214 children in the standardization sample show goodness of fit to the Rasch model (acceptable standard = 95%, sample not described but includes 2 024 children with neurological developmental disorders including spina bifida, multiple unspecified developmental disorders (e.g. cerebral palsy) (Fisher and Merritt, 2012) .	cerebral palsy (Russo et al., 2007, Bonnier et al., 2006).	improvement of 0.5 on the AMPS logit scale (Russo et al., 2007).
Jebsen–Taylor Test of Hand Function (JTT) No evidence supports the JTT for use in the evaluation of upper limb Activity with children with cerebral palsy. Norm-referenced.	Six years onwards.	Not reported.	Not reported.	No psychometric reports known for children with hemiplegia, either for the JTT or any other tests of speed and dexterity (Gilmore et al., 2010). Not indicated as valid or reliable for children with cerebral palsy (Lemmens et al., 2012). SHUE tested for concurrent validity using JTT, a purpose for which it is not accepted as a standard criterion measure (Gilmore et al.,	Not responsive in one trial (Gordon et al., 2007) but did show response in another trial (writing task omitted) (Gordon et al., 2006)	Not reported.

				2010).				
The Functional Independence Measure for Children (WeeFIM) Not valid for use with children with cerebral palsy due to high ceiling effects and poor responsiveness (Debuse and Brace, 2011) Not recommended for evaluation of functional change after orthopaedic surgery in children with unilateral cerebral palsy (Sanders et al., 2006) Designed primarily to measure self-care, not upper limb Activity (Gilmore et al., 2010) Validated in children with cerebral palsy for presenting functional status rather than as an outcome measure (Azaula et al., 2000)	Children aged six months to eight years.	Not reported.	Severe ceiling effect (12% to 56%) in moderate to mildly affected children with cerebral palsy (Debuse and Brace, 2011).	ICC = 0.97 (Ottenbacher et al., 1996).	ICC 0.73–0.97(Ottenbacher et al., 1997).	Needs further validity testing for children with cerebral palsy (Harvey et al., 2008).	Mixed reports of responsiveness (Harvey et al., 2008). Poor responsiveness (Debuse and Brace, 2011). Responsiveness reported in general population of disabled children only over a period of one year (Ottenbacher et al., 2000)	Not reported.
Shriners Hospital Upper Extremity Evaluation (SHUEE) Validated only for American	American children with	Not reported.	Not reported.	Based on spontaneous functional	Based on spontaneous functional	Fair correlation with self-care of PEDI (Pearson	Responsiveness not reported (Lemmens et al., 2012) but there	Not reported.

children with unilateral involvement (Gilmore et al., 2010). Validation as part of assessment of surgical outcomes in children with cerebral palsy (Davids et al., 2006) and using an inappropriate validation techniques (Gilmore et al., 2010).	hemiplegic CP aged three to eighteen years (Gilmore et al., 2010, Klingels et al., 2010).			element (Pearson product-moment correlation coefficient = 0.99) (Davids et al., 2006).	element (Pearson product-moment correlation coefficient = 0.90) (Davids et al., 2006).	product-moment correlation coefficient =0.47)(Davids et al., 2006) SHUE tested for concurrent validity using JTT, for which it is not accepted as a standard criterion measure (Gilmore et al., 2010).	were changes in spontaneous function and dynamic positioning ($p < 0.0001$ for each) but not grasp and release following surgery (Smitherman et al., 2011).	
House Functional Classification (HFC) including the modified HFC (mHFC) HFC originally developed for assessment of surgical outcomes (for TIP) on children with cerebral palsy, using arbitrarily chosen scoring categories and with no psychometric testing (House et al., 1981). Further development and reliability	Children with cerebral palsy aged Three to eighteen years.	Not reported.	Not reported.	HFC ICC=0.92; mHFC ICC=0.94 (Koman et al., 2008).	HFC ICC=0.94; mHFC ICC=0.96 (Koman et al., 2008).	Concurrent validity with MAUULF = 0.84 (Koman et al., 2008).	mHFC unresponsive in an RCT looking at functional benefits of botulinum toxin used in treatment of spasticity of children with cerebral palsy (Koman et al., 2013). In the same trial, the (MAUULF) did show an improvement.	Not reported.

testing carried out by Konan et al (Koman et al., 2008).								
Pediatric Outcomes Data Collection Instrument (PODCI) A generic instrument (Barnes et al., 2008). The PODCI was designed for assessment of surgical outcomes (Debuse and Brace, 2011, Harvey et al., 2008). Based on psychometric evidence, PODCI ruled out as valid measure of function for cerebral palsy population (Debuse and Brace, 2011) Only the Sports and Physical Function domain fits the Rasch measurement model in children with cerebral palsy (Seok et al., 2012) .	Two to eighteen years: there is a Parent version for children aged four to eleven years, and a version for children aged eleven years and over (Barnes et al., 2008).	Upper extremity function domain showed floor effects in a Rasch analysis of children likely to have no severe upper limb problems (Seok et al., 2012) and in 11% of a small sample of children with cerebral palsy, 42% of which were classified as lower limb involvement only (McCarthy et al., 2002).	Upper extremity function domain showed ceiling effects up to 43% in a small sample of children with cerebral palsy (McCarthy et al., 2002).	Pearson correlation coefficient = 0.94 and 0.96 for parents and child questionnaires' Upper Extremity domain respectively (Daltroy et al., 1998).	Not reported.	Correlation of Upper Extremity Function with physician rating = 0.62 (Daltroy et al., 1998). Internal consistency Cronbach's alpha 0.76–0.97 (Harvey et al., 2008). Concurrent validity with self-care of PEDI Pearson product-moment correlation coefficient = 0.85 (McCarthy et al., 2002). Barnes et al (Barnes et al., 2008) advise caution with the validity of the	Responsiveness needs to be established for children with CP (Harvey et al., 2008).	Not reported.

						child (four to eleven years) PODCI in ambulatory children.		
Nine hole peg test (NHPT) and the Box and Block Test (BBT) No evidence of appropriate validity for the use of these measures in children with cerebral palsy, and some evidence that they do not relate to daily Activity using the upper limb.	The only validation studies carried out in children with cerebral palsy aged seven years six months to thirteen years and four months old (Van Hedel and Wick, 2011).	N/A	N/A	There is limited support for the use of the NHPT and the BBT for use with children with CP or for evaluation of Activity Limitation. A conference oral presentation (Van Hedel and Wick, 2011) suggests that they do not relate to use of upper limb in daily life. The BBT may be more suitable for screening children who may or may not have problems with fine motor dexterity (Smith et al., 2000).			Not reported.	Not reported.
Video Observations Aarts and Aarts (VOAA) Measures actual performance and level of developmental disregard of child engaged in two	Children with unilateral cerebral palsy aged two years six months	Not reported.	Not reported.	ICC 0.87–1.00 (Klingels et al., 2010).	kappa 0.62–0.67 (Klingels et al., 2010). ICC = 0.95–1.00 (Klingels et al., 2010).	Construct validity assessed by comparing results of non-disabled children with children with	Not investigated but may be unresponsive to some changes in performance (Houwink et al., 2013) .	Not reported.

standardised tasks (Houwink et al., 2013). Tasks evaluate reaching, grasping and holding (Aarts et al., 2007) but the developers propose that is a discriminative tool rather than an evaluative measure (Aarts et al., 2009).	to ten years (Aarts et al., 2009, Aarts et al., 2007).					cerebral palsy. Maximal ability (capacity) was lower in children with cerebral palsy but did not quite reach statistical significance (p = 0.052). Actual ability shown by the children (performance) was lower in children with cerebral palsy (p <0.001).		
Assessment of Children's Hand Skills (ACHS) related to the CHSQ (see below) A generic instrument, psychometrically tested using Rasch analysis, validated in Australian children (Chien et al., 2010) and Taiwanese children (Chien et al., 2011a,	Developed on a generic population of Taiwanese children aged Two to twelve years old	Not reported.	Not reported.	kappa ≤ 0.79.	Inter-rater agreement Spearman's rho = 0.63 (Chien et al., 2010). Inconsistent and unsatisfactor	Content validity: established through the literature, use of other hand measures and expert opinion (Chien et al., 2011b). Some children with	Not reported.	Not reported.

Chien et al., 2011b). Needs further psychometric testing. Bilateral hand use tested (Chien et al., 2011b).	(Chien et al., 2011b).		y inter-rater agreement in some items (Chien et al., 2010).	cerebral palsy did not fit the statistical expectations of the Rasch model and were removed from the testing (Chien et al., 2011b). Internal consistency not evaluated (Chien et al., 2010).
Children's Hand-Skills ability Questionnaire (CHSQ) related to the ACHS as a companion assessment A generic instrument, validated only in Australia and in Taiwanese children and psychometrically tested using Rasch analysis (Chien and Brown, 2012), it needs further development and	Developed on a generic population of Taiwanese and Australian children aged two to twelve	Developed using Rasch analysis, the authors report further development is still necessary (Chien and Brown, 2012).		

psychometric testing.	years old (Chien and Brown, 2012).		
Children's Hand-use Experience Questionnaire (CHEQ) A questionnaire focusing on activities requiring bilateral hand use, psychometrically tested using Rasch analysis (Skold et al., 2011). There is no evidence that the questionnaire was re-tested after its subsequent modifications following the Rasch analysis. Validated only in Sweden.	Validated in Sweden children aged six to eighteen years for use with unilaterally affected disability including unilaterally-impaired children with cerebral palsy aged eight to fourteen years old.		The CHEQ was developed using a sound strategy and then psychometrically tested using Rasch analysis (Skold et al., 2011). The resulting modifications have not been subjected to psychometric testing. The sample size is very small, adequate for a pilot testing but the item calibrations and person measure estimates will have wide confidence intervals for their values. Some items appear inappropriate for all children e.g. fastening a necklace (12.5% of children report some activities did not apply to them (Skold et al., 2011).

COPM : Canadian Occupational Performance Measure; ICF-CY: International Classification on Function, Disability and Health for Children and Youth; BOTMP: Bruininks-Oseretsky Test of Motor Proficiency; GAS: Goal Attainment Scale; PMAL: Paediatric Motor Activity Log; PEDI: Paediatric Evaluation of Disability Inventory; PDMS; Peabody Developmental Motor Scales ; MAUULF: Melbourne Assessment of Unilateral Upper Limb Function; QUEST: Quality of Upper Extremity Skills Test; AHA: Assisting Hand Assessment; AMPS: Assessment of Motor and Process Skills; JTT: Jebsen–Taylor Test of Hand Function; WeeFIM: The Functional Independence Measure for Children; SHUEE: Shriners Hospital Upper Extremity Evaluation; ADL: Activities of Daily Living: HFC: House Functional Classification; mHFC: modified

House Functional Classification; PODCI: Pediatric Outcomes Data Collection Instrument; NHPT: Nine hole peg test; BBT: Box and Block test; VOAA: Video Observations Aarts and Aarts; ACHS: Assessment of Children’s Hand Skills; CHSQ: Children's Hand-Skills ability Questionnaire; CHEQ: Children’s Hand-use Experience Questionnaire; MCID: Minimal Clinically Important Change; SRM: Standardized Response Mean; ICC: intraclass correlation coefficient; CP: cerebral palsy.

Gilmore et al. (2010) carried out a good quality systematic review on upper limb activity measures for children aged 5 to 16 years old with unilateral impairment. They rejected the Bruininks-Oseretsky Test of Motor Proficiency (BOTMP), the Jebsen–Taylor Test of Hand Function (JTT), Box and Block Test (BBT) and Nine-Hole Peg Test (NHPT), the Paediatric Motor Activity Log (PMAL) and Peabody Developmental Motor Scales (PDMS) because there was no evidence for validity and reliability in this population. This search revealed no further evidence published in the three years since the review by Gilmore et al. (2010).

The PMAL has undergone Rasch analysis (Wallen et al., 2009a, Wallen et al., 2009b). This resulted in modifications to the PMAL that still require further psychometric testing (Wallen et al., 2009a, Gilmore et al., 2010). The norm-referenced BOTMP (Bruininks and Bruininks, 2005) and the norm-referenced JTT (Klingels et al., 2010) are both used in trials to evaluate interventions to improve upper limb function in children with cerebral palsy (Hoare and Imms, 2004, Sakzewski, 2012, Gordon et al., 2006), which is why they are still included in this appraisal.

The weeFim is also in common use to evaluate trial outcomes (Hoare et al., 2007b). It is a generic measure that Debuse and Brace (2011) suggest should not be used with children with cerebral palsy.

The Children's Hand-use Experience Questionnaire (CHEQ) has been developed using a sound strategy and evaluated using Rasch analysis but the resulting modifications have undergone further psychometric testing and some items appear unsuitable for all children e.g. fastening a necklace (Skold et al., 2009, Skold et al., 2011). Similarly, the authors of the Children's Hand-Skills ability Questionnaire (CHSQ) report that development and psychometric testing is incomplete (Chien and Brown, 2012). The Peabody Developmental Gross Motor Scale (PDMS) is also norm-referenced (Palisano et al., 1995); it should be used for discriminative purposes only and shows poor responsiveness (Palisano et al., 1995). Limited evidence supports the use of the Nine Hole Peg Test (NHPT) and the Box and Block Test (BBT) as screening tools but not as evaluation tools (Smith et al., 2000). Van Hedel and Wick (2011) suggest that they are not at all associated with upper limb activity.

The Pediatric Outcomes Data Collection Instrument (PODCI) and the Assessment of Motor and Process Skills (AMPS) are generic instruments (Barnes et al., 2008, Fisher and Merritt, 2012). The AMPS is unresponsive in trials (Bonnier et al., 2006, Russo et al., 2007), and the POCDI has no evidence for its responsiveness (Harvey et al., 2008). The AMPS and the Paediatric Evaluation of Disability Inventory (PEDI) are norm-referenced, which suggests they are suitable only as discriminative tools (Palisano et al., 1995, Rosenbaum et al., 1990). A small validation study for children with cerebral palsy has been carried out for the PEDI but the Rasch analysis used data from non-disabled children (Nichols and Case-Smith, 1996). The PEDI was unresponsiveness in a clinical trial (Russo et al., 2007); other studies report responsiveness as inconclusive (Harvey et al., 2008). It is reported as requiring further validity testing for children with cerebral palsy (Harvey et al., 2008) so it is not appropriate for use in this patient population or for this health condition.

The PODCI and the HFC were developed for evaluating surgical outcomes, the latter particularly for thumb-in-palm deformity (Debuse and Brace, 2011, Harvey et al., 2008, House et al., 1981). There is no evidence that they are responsive in the evaluation of activity limitation (Harvey et al., 2008, Koman et al., 2013) and the PODCI has no evidence of its validity for use in children with cerebral palsy (Debuse and Brace, 2011). The PODCI and PEDI were excluded from the systematic review by Gilmore et al. (2010) because they contain a large proportion of items that do not relate to upper limb activity (online supplemental table in Gilmore et al., 2010), suggesting unacceptable multi-dimensionality.

Of the 21 measures, only three were found to have acceptable psychometric standards for evaluating upper limb activity of children with cerebral palsy: the Assisting Hand Assessment (AHA) (Gilmore et al., 2010, Greaves et al., 2010, Klingels et al., 2010), the ABILHAND-kids (Gilmore et al., 2010, Klingels et al., 2010) and the Melbourne Assessment of Unilateral Upper Limb Function (MAUULF) (Gilmore et al., 2010, Klingels et al., 2010). The AHA is a Rasch-derived measure with good quality evidence for its psychometric standards (Krumlinde-Sundholm et al., 2007, Eliasson et al., 2005); the ABILHAND-kids is also Rasch-derived with good psychometric standards (Arnould et al., 2004) but there are suggestions that the adult version, which is scored identically to the ABILHAND-KIDS, lacks responsiveness (Bovolenta, 2009). The ABILHAND-kids was evaluated using children with

cerebral palsy at the upper spectrum of functional ability which makes it poorly validated for the evaluation of children at the lowest levels of functioning (Arnould et al., 2004), this increases the likelihood of poor responsiveness at these lower levels. Although the MAUULF and the Quality of Upper Extremity Skills Test (QUEST) are widely used in clinical trials, they do not focus on upper limb activity (Hoare et al., 2011), with a large number of items relating to body functions and not activity, which suggests a lack of unidimensionality. They are unresponsive in clinical trials (Wallen et al., 2007). A Rasch analysis suggests that 25% of MAUULF items have poor psychometric properties that need to be addressed (Randall et al., 2010).

Goal Attainment Scaling (GAS) and the Canadian Occupational Performance Measure (COPM) were omitted from the systematic review by Gilmore et al. (2010) because it was stated that they were goal-setting tools. However, these measures are for developing individualised goals and their use for evaluating change in activity limitation has been established (Cusick et al., 2007, Steenbeek et al., 2007). The activities comprising the goals are defined in discussions between therapists, families and children (Cusick et al., 2007, Steenbeek et al., 2007, Tennant, 2007, Verkerk et al., 2006). There is good evidence for their responsiveness (Sakzewski et al., 2007) which has been demonstrated in clinical trials (Wallen et al., 2007) but there are reservations about the standard of psychometrics of individualised measures such as these (Tennant, 2007). This includes the use of arithmetic procedures on ordinal data (Tennant, 2007).

Gilmore et al. (2010) conclude that the Shriners Hospital Upper Extremity Evaluation (SHUEE) and the QUEST may be appropriate for evaluating changes in activity limitation for children with unilateral impairment following surgery or a spasticity intervention. However, their paper presents no evidence that the SHUEE has been tested adequately for validity or unidimensionality as, like the QUEST, its outcome score includes aspects of other dimensions (i.e. body functions). The assertion that the outcome scores for both of these measures are percentages is incorrect: these measures produce ordinal outcome scores.

There is no evidence for the use of the Video Observations Aarts and Aarts (VOAA) as a measure to evaluate changes in performance of upper limb activity, but Aarts et al. (2009) describe it as a discriminative tool between children with activity limitations and non-disabled children. According to the developers, the Children's Hand-Skills ability Questionnaire (CHSQ) and the Assessment of Children's Hand Skills (ACHS) require further development and testing before their psychometric standards and their acceptability for use with children with cerebral palsy can be evaluated (Chien and Brown, 2012). Testing and development so far has only been conducted in Australia and Taiwan (Chien et al., 2010, Chien et al., 2011a, Chien et al., 2011b).

3.3.4 Discussion of critical appraisal

Outcome measures must be selected with consideration given to a number of factors e.g. the variable being measured, the purpose of the measurement, the appropriateness of the measure to the target population, psychometric properties, cost, and burden of administration (Wagner and Davids, 2012). The aim of this critical appraisal was to review these characteristics and properties for all outcome measures used in research to evaluate the upper limb activity of children with cerebral palsy. This would provide the information necessary for selecting the most appropriate and scientifically robust outcome measures for this research study.

In considering the measures, the advantages of modern psychometric techniques wre tken into account, the evidence that ordinal outcome scores possess wide confidence intervals and characteristics unsuitable for use in arithmetic and parametric procedures, and the view that a measure should be designed and tested in the population and health condition for which it is intended to be used. Unidimensionality is a fundamental principle of measurement that must be evaluated and found to be acceptable in any measure that purports to measure a variable, trait or characteristic.

The standard of psychometric testing of measures varied widely, as identified in two earlier reviews (Greaves et al., 2010, Wagner and Davids, 2012). Of the 21 measures identified as being used to evaluate upper limb activity limitation in children with cerebral palsy, only two of 21 measures (the AHA and the ABILHAND-kids) possess psychometric qualities robust enough for this purpose; others have undergone little testing or show very poor

psychometric properties for their use with children with cerebral palsy (the JTT, NHPT, BBT). Some measures which showed potential require further psychometric testing and development (MAUULF, PMAL, ACHS/CHSQ and CHEQ).

Measures must be validated on a sample of the intended target population. However, many measures that purport to have established validity do not appear to have been adequately validated in children with cerebral palsy. Criterion validation of a measure can only be performed when the established measure is itself adequately validated for use with children with cerebral palsy. Adequate validation for children with cerebral palsy has not been established with the BOTMP, PDMS, JTT, NHPT, BBT, ACHS/CHSQ, or the CHEQ.

Poor responsiveness is a characteristic of generic measures (Arnould et al., 2012, McCullough and Parkes, 2008). Generic measures included in this review that are in widespread use with children with cerebral palsy are the AMPS, PODCI, ACHS/CHSQ, PEDI, PDMS and WeeFIM. The AMPS was demonstrated as having poor responsiveness and, along with the WeeFim, is not designed to measure activity outside of self-care (its primary purpose) (Gilmore et al., 2010). There is no evidence that the JTT or the WeeFIM are appropriately validated for use with children with cerebral palsy (Debuse and Brace, 2011, Gilmore et al., 2010, Harvey et al., 2008, Lemmens et al., 2012).

The evidence collected from this systematic search and critical appraisal suggests that the PMAL, PODCI, VOAA, ACHS/CHSQ, CHEQ, BBT and NHPT should not be used for the functional evaluation of children with cerebral palsy, but there is potential that the PMAL, ACHS/CHSQ and CHEQ may prove useful measures for this purpose if subjected to further adequate psychometric testing.

Two reviews suggest that the MAUULF should be included alongside the ABILHAND-kids and the AHA as possessing the most robust psychometric properties (Gilmore et al., 2010, Klingels et al., 2010). However, less than half of the MAUULF's items relate to activity limitation, and a post-development Rasch analysis has identified that a quarter of all items require further psychometric testing. A reservation for use of all three measures in the UK is that the MAUULF has only been validated for use with Australian children, the

ABILHAND-kids for French-speaking Belgian children and the AHA has been validated for Swedish children; all three therefore require psychometric evaluation in the UK.

Both the AHA and the ABILHAND-kids were developed using Rasch analysis. The AHA focuses on the impaired upper limb in unilaterally-impaired children. The ABILHAND-kids is a questionnaire containing unilateral and bilateral activities. The AHA focuses on the assistance that a child has from his or her impaired upper limb and assumes that the child uses the non-impaired upper limb for all functional activity. This contrasts with the ABILHAND-kids measure, which measures ease of achievement of both unilateral and bilateral activities so that children with unilateral impairment are likely, naturally, to achieve a higher score than children with bilateral involvement. However, the ABILHAND-kids was validated on a sample with very few children of limited functional ability (Arnould et al., 2004), and doubts have been expressed about its responsiveness (Bovolenta, 2009). The AHA is therefore more likely to identify changes in the affected arm of children with unilateral impairment. The financial cost, lengthy training and difficult scoring associated with the AHA are potential deterrents for its acquisition by rehabilitation researchers and clinicians.

One systematic review (Gilmore et al. (2010)) concludes that the SHUEE and the QUEST may be appropriate for evaluating changes in activity limitation for children with unilateral impairment following surgery or a spasticity intervention, however, there is no evidence that the SHUEE has been tested adequately for validity or unidimensionality as its outcome score includes aspects of other ICF-CY domains (Body Function and Structures) (Gilmore et al., 2010). The QUEST dimensions' ordinal outcome scores are generated using mathematical procedures which are inappropriate for ordinal data. The GAS and COPM have been validated for use with children with cerebral palsy. They appear responsive but there are suggestions that health care professionals (legitimately) generate individualised goals at which the patient is most likely to show improvement. Furthermore, the non-linear nature of the ordinal outcome scores and the inappropriate mathematical procedures used in outcome score generation can exaggerate clinically significant change (Stucki et al., 1996, Tennant, 2007). There are also wide confidence intervals around ordinal outcome scores (Hobart et al., 2007). The individualised nature of GAS and COPM goals mean that these measures are inappropriate for group comparisons in research.

These findings are the result of a systematic search to locate all articles that describe or include details of any measure that is used to evaluate upper limb activity limitation of children with cerebral palsy, and the collation of all details relating to their psychometric testing and use. The major limitation in the conduct of this appraisal is that it did not take the form of a systematic review. There were no exclusion criteria because the intention was to gather as much information about the measures as possible and to establish the quality of the literature on which use of measures was based. Critical appraisal of the literature relied not on quality criteria, but on consideration of advantages and disadvantages of the different approaches to psychometric testing and of the limitations surrounding the use of ordinal data within mathematical operations.

3.3.5 Summary and outcome of the findings of the critical appraisal

3.3.5.1 Summary of findings

Not all measures included in this appraisal are appropriately validated for use with children with cerebral palsy. Many are unresponsive and their high reliability coefficients may be artificially raised because of the high number of items. The AHA and the ABILHAND-kids offer therapists and researchers the most scientifically robust option for accurately and appropriately evaluating activity limitation of children with cerebral palsy, although there is still further psychometric testing to be performed on these measures. There is evidence of floor and ceiling effects with the ABILHAND-kids, and doubts have been expressed about its responsiveness. The AHA is difficult to score; it is costly in terms of training and financial outlay.

3.3.5.2 Outcome of appraisal and action to address limitations

This critical appraisal of the psychometric properties of measures of upper limb activity limitation of children with cerebral palsy supports the suggestions by previous researchers that the psychometric properties essential for measures used in research (and clinical practice) are largely absent or weak. Even measures which have undergone thorough development using Rasch analysis and which the evidence suggests are the best available – the AHA and the ABILHAND-kids - possess weaknesses with their use, scoring or their properties.

Based on the evidence of the critical appraisal described in section 3.3, the ABILHAND-kids was selected as the primary measure for use in this research study. To evaluate individual

change, the COPM was selected for its responsiveness. However, the limitations of these measures is recognised e.g. potential floor effects and suspected lack of responsiveness for the ABILHAND-kids, and the individualised goal setting and nature of ordinal outcome scores of the COPM.

3.3.5.3 Future action based on appraisal findings

Although the critical appraisal identified which measures had acceptable psychometric properties and characteristics for use in this study, it also suggested that the development of a new measure was necessary. As well as testing the new measure for acceptable psychometric standards, the use of Rasch analysis would also identify whether scientific standards permitting conversion of raw scores to interval-level outcome scores had been achieved.

A number of the measures identified in the appraisal included inappropriate items, e.g. age-inappropriate or gender-inappropriate. Item responses were limited to either two or three in some unresponsive measures, suggesting that increasing the number of response categories would improve responsiveness. Some measures included items that assessed characteristics other than activity limitation, while some measures were not appropriate or validated for children with cerebral palsy. This suggested that the new measure would have to be conceptualised, designed and constructed from new items and item responses, specifically for activity limitation of children with cerebral palsy.

The development of this new measure of upper limb activity limitation therefore became the first stage of this research study. By addressing the problems identified in other measures, this new measure should meet all psychometric standards for its intended use, that of evaluating change in upper limb activity limitation of children with cerebral palsy.

3.4 Development of a new measure of upper limb activity limitation for children with cerebral palsy – the Children's Arm Rehabilitation Measure (ChARM)

3.4.1 Introduction

The critical appraisal in the previous section illustrated several psychometric problems common with outcome measures to evaluate upper limb activity limitation in children with cerebral palsy. These included multi-dimensionality, floor and ceiling effects, potentially artificially-increased reliability, a lack of validity, doubtful reliability, and inappropriate

statistical operations on ordinal data. The ordinal-level outcome scores of measures also present a potential source of inaccuracy when drawing conclusions from research, supporting the argument that only interval-level outcome scores are acceptable for research. There is therefore a requirement in research and rehabilitation for children with cerebral palsy for establishing upper limb activity limitation outcome measures that are appropriately validated, reliable, responsive and psychometrically sound.

Diligent design and development of outcome measures from the "bottom up" may help to prevent the limitations described above (Hobart et al., 2007), starting with careful consideration of the actual variable that is to be measured (Hobart et al., 2007, Streiner and Norman, 2003). Even abstract characteristics (e.g. pain, mood, tiredness that are sometimes described as latent traits or constructs) can be measured with appropriate instruments e.g. questionnaires with carefully selected and properly-developed items (Wilson, 2005). However, defining and selecting items that accurately represent or capture the variable which is to be measured is of critical importance (Wilson, 2005). Post-development psychometric testing to establish the measurement properties of a new measure is essential (Hobart et al., 2007, Tennant, 2007) but nothing in the subsequent validation of a measure can rectify vague or inappropriate items (Streiner and Norman, 2003).

Response options for items also need to be properly developed. Item responses can be structured with a variety of types and number (Streiner and Norman, 2003), or there is the option of the same number and choice of response options for each item, for example rating capability as 'easy/difficult/impossible' for each item, as in the ABILHAND-kids. Other measures include a greater number of response options that rate level of achievement, such as the weeFIM which has seven response categories. Too many response options can introduce error; for example, two raters might agree on the level of the variable being assessed, but tick adjacent responses from the choice of several, reducing reliability of the measure (Bond, 2003). Conversely, too few response options may result in poor responsiveness (Bovolenta, 2009) possibly as a consequence of increased floor and ceiling effects. Bond and Fox (2001) suggest that the optimum number of response options is entirely dependent on the variable being measured and should be assessed empirically for each scale. Testing of item responses has been improved through

the use of modern psychometric techniques such as Rasch analysis (Tennant and Conaghan, 2007).

The aim of this stage of the study is to develop a questionnaire designed to measure upper limb activity limitation of children with cerebral palsy: the Children's Arm Rehabilitation Measure (ChARM). This sub-study has two parts. Part 1 has a number of objectives:

1. using a novel technique to produce valid, appropriate items and response options valid for children with cerebral palsy;
2. to improve face validity;
3. to finalise the structure and style of the items, response options and design of the ChARM questionnaire.

Part 2 of the ChARM sub-study will evaluate the ChARM's psychometric properties by applying the Rasch model to a dataset of responses from a large sample of children with cerebral palsy. Items and responses will be modified based on the results of the Rasch analysis, and the psychometric properties re-assessed on a second dataset of responses. The aim is to produce a measure of upper limb activity limitation that is validated for children with cerebral palsy aged between five and sixteen years old, and that is unidimensional, responsive to changes in activity limitation, will overcome the problem of ceiling or floor effects and will allow the transformation of the ordinal raw scores into linear outcome scores.

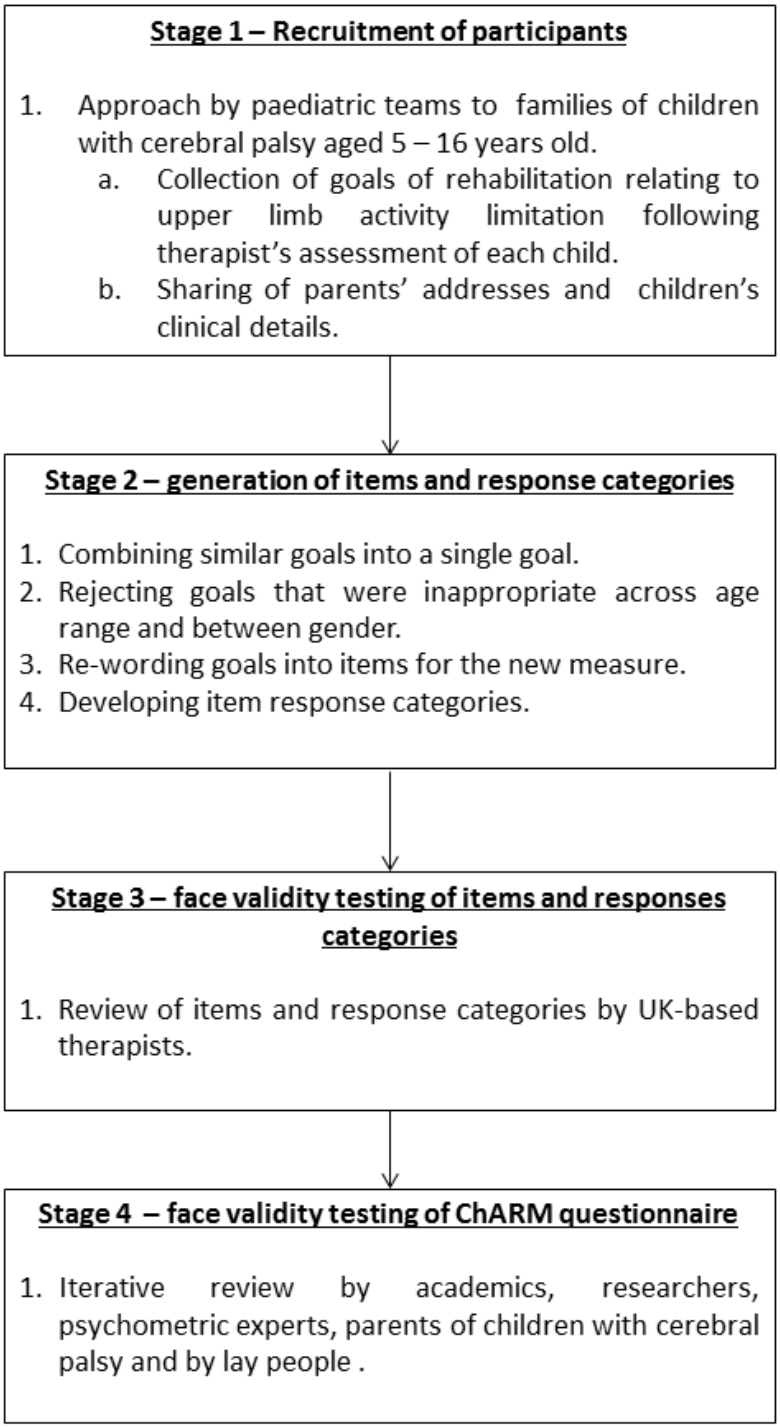
3.4.2 Methodology for development of the ChARM questionnaire

In addition to the use of outcome measures when evaluating the child's potential for activity and participation, therapists (and researchers) undertake detailed assessments to determine the child's goals of treatment (or research outcomes). These assessments usually include discussions with parents and children to identify obstacles that limit children's activity and restrict their participation in life situations. The Canadian Occupational Performance Measure (COPM) is well suited to this aspect of rehabilitation assessment. Initial development of the questionnaire items was based on the hypothesis that rehabilitation goals formed from functional assessment of children with cerebral palsy aged five to sixteen years old by medical staff and therapists involved in clinical practice and research would provide a valid basis for items which relate directly to upper limb activity of children with cerebral palsy. The inclusion criteria were therefore restricted to children with cerebral palsy aged five to sixteen years old, with no exclusion criteria.

When working on activity-based goals of therapy, therapists break down the activity into natural stages of achievement, each of which provides a natural short-term goal. These stages were hypothesised to provide natural categories for item responses because they are well-defined, ordered categories of ability easily identified by parents trying to categorise the level of their child's activity limitation.

It was recognised that the opinion of experienced clinicians and parents of children with cerebral palsy would be essential to improve and finalise acceptable descriptions and clarity of the responses and response options; it was also recognised that members of the public unconnected with research and health care could make an important contribution in the production and face validity testing of the questionnaire. Development of the ChARM questionnaire was therefore divided into 4 stages, shown in Figure 3-4 below:

Figure 3-4. Stages of sub-study to develop the ChARM.



3.4.2.1 Research governance approval (ethical and NHS favourable opinion)

Ethical favourable opinion was obtained from the East Yorkshire and North Lincolnshire Research Ethics Committee (REC Reference 10/H1304/46) on 28th September 2010. NHS permission was obtained from each of the UK's National Health Service (NHS) institutions

in which the paediatric teams were employed. The study was registered on the National Institute for Health Research (NIHR) Clinical Research Network (CRN) portfolio (ID 9600).

3.4.2.2 Stage 1: recruitment of participants

In order to elicit the necessary support of therapists for the purpose of obtaining goals of rehabilitation the study was presented to the ten paediatric therapy teams in Table 3-6. Only one team refused to support the study, citing excessive workload and staff pressures as the reason, therefore nine paediatric teams supported this stage of the sub-study.

Table 3-6. UK paediatric teams' sites that supported Stage 1.

Airedale NHS Foundation Trust	
Calderdale and Huddersfield NHS Foundation Trust	
Leeds Community Healthcare (formerly Leeds Primary Care Trust)	
Bolton NHS Trust	
Harrogate and District NHS Foundation Trust	
Sheffield Children's NHS Trust	
Mid Yorkshire NHS Trust	
York NHS Trust	
Tameside and Glossop PCT	
Bradford Teaching Hospitals NHS Foundation Trust	*Refused*

Goals of therapy were to be collected prospectively. This was to reduce the burden associated with supporting the study on therapy teams – the functional assessments would be part of their clinical work and completing a Case Report Form would take only a few minutes, whereas searching through case files to retrospectively identify children and appropriate goals would be a resource-intensive task which had the potential to prevent therapy teams' participation in the study. When arranging appointments with the child and their parents, therapists would introduce the study using the text in Table 3-7 below. Along with the appointment confirmation, therapists would post out a study package containing Participant Information Sheets and Informed Consent (PISIC) forms, and a stamped envelope pre-addressed to the study Chief Investigator for the return of the Informed Consent form. If no appointment was posted out, therapists would give the parents and child the study package at the appointment. Obtaining informed consent was necessary to

allow the sharing of identifiable information such as addresses, required for posting the developed ChARM questionnaire to parents in preparation for Rasch analysis. Therapists were permitted to collect the completed informed consent form themselves at the appointment but were not permitted to obtain informed consent; any questions or requests for further information were to be directed to the Chief Investigator.

Table 3-7. Guide for therapists introducing the study to potential participants.

"We are supporting a study by the University of Leeds to develop a new instrument for measuring how well children with cerebral palsy use their arms and hands. I will send you an introductory letter, consent forms and information sheets about the study along with your appointment letter. If you have any questions or you would like more information, please contact Nick Preston, the lead researcher, who will call you straight back. Nick's contact details are on the letters and information sheets. If you agree to help us, please make sure you complete the consent form before the appointment and give it to me at the appointment. If you do decide to take part, you can change your mind at any time, and it won't affect your care or treatment by the NHS".

Once informed consent was received by the Chief Investigator the therapists were permitted to share personal and clinical details of the children, including goals of rehabilitation relating to activity limitation of the upper limb.

In order to power the Rasch analysis it was necessary to recruit a minimum of 150 participants. Given the wide area and size of the population covered by the paediatric teams, no problems were anticipated with achieving this target within 12 months. However, this process gained the support of only 19 families in the first 10 months of the study. One reason identified for this poor response was that a number of therapists were judging families' suitability to take part, selecting for inclusion only those families thought by the therapists as likely to participate. Secondly, the detailed PISIC and request to complete and post an Informed Consent form may have deterred interest and participation. A final possible reason is the impact on paediatric team resources of the austerity measures and financial pressures paced upon institutions by the UK government during this period. One team stated that these pressures were behind their decision not to support this sub-study; another participating team reported that for a large proportion of the sub-study lifetime they were unable to see any families or children for face-to-face contacts or commit any time to the sub-study because the team was reduced to only two members of staff.

To increase the number of goals received, and to ensure that a sufficient quantity of goals was received from the complete age range and range of disability (assessed using the Manual Ability Classification System (MACS; Eliasson et al., 2006), a protocol amendment was submitted to the Research Ethics Committee (REC). The amendment described a new procedure that collected retrospective anonymised goals of therapy from the therapy notes of all children with cerebral palsy known to the supporting teams' therapists. This approach was avoided in the initial protocol because of the additional burden on therapy teams' resources but teams were supportive of this method to generate further goals. Collection of retrospective goals of rehabilitation did not affect the validity of the items, because assessment of upper limb activity limitation of children with cerebral palsy is independent of this study's aims and methodology.

For the purposes of achieving enough questionnaires for the Rasch analysis, the amendment proposed to provide a large quantity of the finalised ChARM questionnaire to each participating paediatric team. Teams would post out the ChARM questionnaire, a clinical details form and a pre-paid pre-addressed envelope to the parents of every child with cerebral palsy aged between 5 years and 16 years old in each NHS institution. Contact details for the Chief Investigator were clearly displayed on all documents, with an invitation to contact the Chief Investigator for more information. Completing and posting the questionnaire to the Chief Investigator implied consent and overcame the requirement for a signed Informed Consent form.

A protocol amendment was also submitted for REC approval for two other changes designed to increase the number of appropriate goals:

a. Activity rehabilitation goals and outcomes reported for those children treated for upper limb spasticity from an audit of 237 children with cerebral palsy attending the Leeds Teaching Hospitals NHS Trust spasticity clinic, from which there were 97 upper limb treatments (Preston et al., 2011), were used.

b. Upper limb activity rehabilitation goals taken from assessments of children taking part in two NIHR-funded research studies to develop and trial assistive gaming technology (National Institute for Health Research, 2013b, Weightman et al., 2011).

3.4.2.3 Stage 2 – generation of items and response categories

Rehabilitation goals from all sources were re-written into item stems for inclusion on the questionnaire. The item stems formed the question to which parents would select the response category that best described the child's achievement of the item. Some goals were rejected as being not appropriate for all children for whom the ChARM is targeted e.g. the goal "to be able to peel vegetables using a knife" is inappropriate for younger children old for safety reasons. Goals that were similar to each other were rephrased to produce a single item stem appropriate for both genders and all ages from 5 to 16 years old. So, for example, "to be able to put on a pullover/jumper" and "to be able to put on a T-shirt" formed the item stem "can your child put on a vest (or short-sleeved T-shirt) if it is laid out properly for them?".

Item responses were formed by characterising the activity of each item into natural stages of achievement of the activity. For example, putting on a sock is taught by therapists in the stages of achievement shown in Table 3-8 below:

Table 3-8. Item responses relating to natural stages of putting on a sock

Item	**Can your child put on their socks by themselves?**
Response 1	Yes, my child can put on their socks by themselves
Response 2	My child can pull a sock on to the toes and up their leg but needs help to tidy up the sock
Response 3	My child can only pull a sock partly onto their toes or foot, and needs help to complete pulling socks all the way up the leg

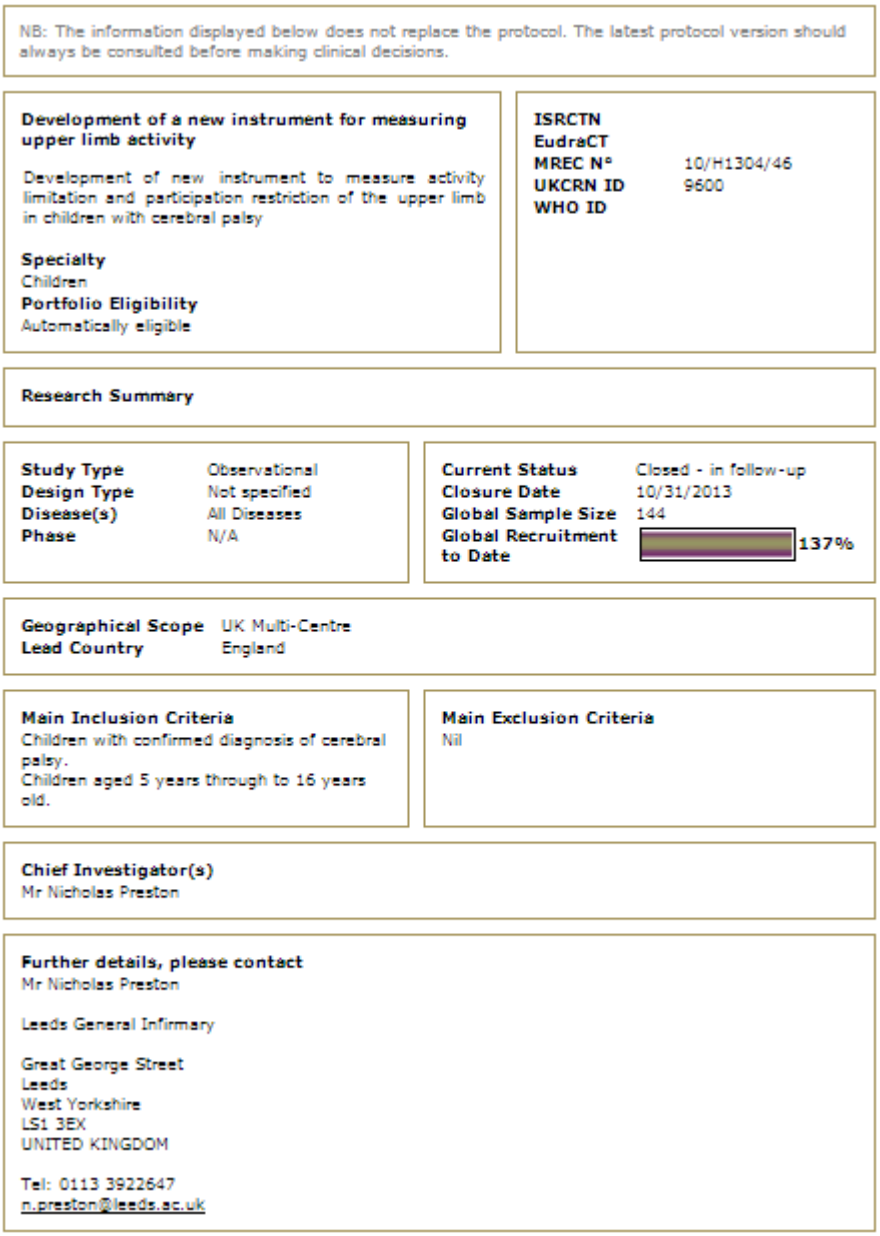
Response 4	No, my child cannot put their socks on at all without help

Each item therefore had a different number of responses, each relating to a natural level of achievement of the activity.

3.4.2.4 Stage 3 – face validity testing of items and responses categories

Adoption of this sub-study onto the portfolio supported not only the process of obtaining NHS permission at each site but placed the study in the public domain through its listing on the CRN Study Portfolio (see Figure 3-5).

Figure 3-5. Listing of ChARM sub-study on the NIHR CRN portfolio.



NB: The information displayed below does not replace the protocol. The latest protocol version should always be consulted before making clinical decisions.

Development of a new instrument for measuring upper limb activity

Development of new instrument to measure activity limitation and participation restriction of the upper limb in children with cerebral palsy

Specialty
Children
Portfolio Eligibility
Automatically eligible

ISRCTN
EudraCT
MREC N° 10/H1304/46
UKCRN ID 9600
WHO ID

Research Summary

Study Type Observational
Design Type Not specified
Disease(s) All Diseases
Phase N/A

Current Status Closed - in follow-up
Closure Date 10/31/2013
Global Sample Size 144
Global Recruitment to Date 137%

Geographical Scope UK Multi-Centre
Lead Country England

Main Inclusion Criteria
Children with confirmed diagnosis of cerebral palsy.
Children aged 5 years through to 16 years old.

Main Exclusion Criteria
Nil

Chief Investigator(s)
Mr Nicholas Preston

Further details, please contact
Mr Nicholas Preston

Leeds General Infirmary

Great George Street
Leeds
West Yorkshire
LS1 3EX
UNITED KINGDOM

Tel: 0113 3922647
n.preston@leeds.ac.uk

This brought the ChARM sub-study to the attention of regional NIHR CRN staff. CRN research facilitators were then able to facilitate and support participation of the three additional paediatric teams shown in Table 3-9 (below) from Stage 3 onwards. Inclusion of these sites and the support of the CRNs proved crucial to the success of the development of the ChARM questionnaire.

Table 3-9. Additional sites that supported the ChARM development.

Cambridgeshire Community Services NHS Trust
Lancashire NHS Trust
Leicester NHS Partnership

The items and response options were reviewed by between two and five therapists in each of the 12 participating teams. Each item was reviewed by at least one therapist who had not developed the rehabilitation goal from which the item was developed. The review resulted in the revision of a number of items and item responses. The item set was then formulated into the ChARM questionnaire.

Stage 3 resulted in 40 items, each with a varying number of response categories. Item stems and response categories had been reviewed and approved by between two and five therapists experienced in the functional assessment and rehabilitation of children with cerebral palsy.

3.4.2.5 Stage 4: face validity testing of ChARM questionnaire

Face validity testing of the ChARM questionnaire was carried out using an iterative approach with paediatric therapists, parents of children with cerebral palsy, academics experienced in the development of new measures, and lay people (see Table 3-10 below). The academics were involved for their expertise in reviewing and critiquing measures in the early stages of development prior to psychometric testing. Paediatric therapists were not from teams that had been involved in the generation of goals or the review of the items. Parents of children with cerebral palsy were an essential component as the end-users of this parent-reported measure, but parents were included who had no experience of disability or academic research to prevent the inclusion of misleading or technical terms, specialist terminology and jargon that is familiar to academics and health care professionals.

After the ChARM was reviewed by the first series of face validity testers and their comments implemented appropriately, the ChARM was presented to a second series of validity testers. This process was repeated two more times and resulted in the correction

of typographical errors, changes to aid clarity and to prevent possible influencing of answers, rewording of confusing terms and sentences, and uniformity of wording of each item and each item's responses. This concluded the development of the ChARM questionnaire and its preparation for psychometric evaluation.

Table 3-10. Iterative process and participants for face validity testing of ChARM questionnaire.

Iteration 1	Iteration 2	Iteration 3	Iteration 4
Academic supervisor, Professor and Consultant of Rehabilitation Medicine with experience of developing new measures and assessing functional outcomes of children with CP	Lay person (parent of healthy child)	Lay person (parent of a child with CP)	Academic supervisor, Professor and Consultant of Rehabilitation Medicine with experience of developing new measures and assessing functional outcomes of children with CP (as iteration 1)
Psychometric researcher specialising in development and evaluation of measures	Lay person (parent of 3 healthy children)	Specialist paediatric occupational therapist	Psychometric researcher specialising in development and evaluation of measures (as iteration 1)
Professor of Movement Cognition experienced at working with children with Development and Coordination Disorder	Lay person (grandparent of 3 healthy children and 8 grandchildren)	Specialist paediatric physiotherapist	Professor of Epidemiology and specialist in psychometrics
PhD psychology student working with children with CP that involves assessing Activity changes	Academic secretary with experience in constructing questionnaires and measures	Specialist paediatric occupational therapist	Specialist Consultant paediatric physiotherapist and Bobath tutor
Chartered Physiotherapist working in research and stroke rehabilitation	Lay person (parent of a child with CP)	Lay person (parent of healthy child)	
PhD student and engineer developing assistive technology for children with CP	Paediatric Occupational Therapist , Team Leader and paediatric spasticity clinic specialist		

3.4.3 Results

Altogether a total of 158 goals, including duplicated goals, were collected from 53 children (34 males, 19 females) with cerebral palsy aged 5 to 16 years old (median 8 years, range 5 – 16 years; mean 8.8 years, SD 2.99 years, MACS levels I to V). The goals collected from these children were mapped to their appropriate categories on the International Classification of Function, Health and Disability for Children and Youth (ICF-CY).

With duplicates removed, there were 78 unique goals from which 40 items were developed for the ChARM questionnaire. The goals are given in Table 3-11, along with their ICF-CY category and the source of the goal e.g. generated by therapy assessment, from the audit of the spasticity clinic or from previous research studies using activity limitation outcomes.

Table 3-11. The 78 unique goals, mapped to ICF-CY categories.

Occupational Therapist-generated goals		
Goal	**ICF-CY code**	**ICF-CY category description**
To be able to maintain upright posture to be able to use a pencil	d4153	Maintaining a sitting position
	d415	Maintaining a body position
To increase functional skills through seating provision	d155	Acquiring skills
To ride a bike without stabilisers	d4402	Manipulating
	d4750	Driving human-powered transportation
To stir cooking pans	d2204	Completing multiple tasks independently
	d6300	Preparing simple meals
	d6301	Preparing complex meals
To tie laces	d4402	Manipulating
To use communication aids or communicate using arms and hands	d340	Producing messages in formal sign language
To access classroom setting with aim of participating in school	d8200	Moving into educational programme or across levels

activity	d8201	Maintaining educational programme
To be able to communicate using picture exchange.	d335	Producing nonverbal messages
To be able to independently dress top half of the body	d540	Dressing
To be able to independently transfer into a supportive chair.	d420	Transferring oneself
To be able to undertake toilet training	d530	Toileting
To help in transfers	d4201	Transferring oneself while lying
To improve transfers into family vehicle.	d4200	Transferring oneself while sitting
To increase speed of hand writing.	d170	Writing
To peel vegetables	d4402	Manipulating
	d6301	Preparing complex meals
Brushing teeth independently	d1550	Acquiring basic skills
	d4453	Turning or twisting the hands or arms
	d5201	Caring for teeth
Drawing diagrams with a ruler independently	d3352	Producing drawings and photographs
Eating with a knife and fork independently	d1550	Acquiring basic skills
	d550	Eating

	d4402	Manipulating
Opening a jar with a powered jar opener	d1550	Acquiring basic skills
	d4401	Grasping
	d4402	Manipulating
	d6300	Preparing simple meals
To be able to document school work using alternative writing aids or a computer	d2105	Completing a complex task
	d3601	Using writing machines
To be able to feed self with fork in left hand	d1550	Acquiring basic skills
	d550	Eating
To be able to handle and count money up to £5.00	d1201	Touching
	d131	Learning through actions with objects
	d2102	Undertaking a single task independently
	d4402	Manipulating
To be able to maintain upright posture to be able to use a pencil	d1450	Acquiring skills to use writing implements
	d1550	Acquiring basic skills

To be able to pour cereal	d4401	Grasping
	d4403	Releasing
	d550	Eating
	d6300	Preparing simple meals
To be able to pour from a jug with a chunky handle.	d4401	Grasping
	d4403	Releasing
	d560	Drinking
To be able to pull trousers and pants up and down	d4403	Releasing
	d4401	Grasping
	d540	Dressing
To be able to put on and remove a vest/shirt/pullover	d540	Dressing
To be able to put on socks using an aid	d2104	Completing a simple task
	d5402	Putting on footwear
To be able to undress top and bottom half with minimal supervision and assistance.	d540	Dressing
	d4402	Manipulating

To be able to self feed using a fork	d550	Eating
To be able to zip/unzip a coat	d4402	Manipulating
To be able to write more	d1450	Acquiring skills to use writing implements
	d440	Fine hand use
To improve at catching balls		
	d4454	Throwing
	d4403	Releasing
	d9201	Sports
To improve at fastening buttons	d4402	Manipulating
To improve transfers from floor to seat	d410	Changing basic body position
To participate in swimming sessions	d4554	Swimming
	d9201	Sports

To self-feed using aids	d550	Eating
To use communication aids or communicate using arms and hands	d335	Producing nonverbal messages
	e1251	Assistive products and technology for communication
	d350	Conversation
To wash self with a sponge	d4401	Grasping
	d510	Washing oneself
	d5100	Washing body parts
	d5101	Washing whole body
Using an easy-grip knife to spread butter on bread	d4402	Manipulating
	d6300	Preparing simple meals
Will write 4 sentences in class with less pain in hand	d1450	Acquiring skills to use writing implements
Cut bread in half using an easy grip knife	d6300	Preparing simple meals
	d1550	Acquiring basic skills
	d4402	Manipulating

Research study outcome goals

Goal	ICF-CY code	ICF-CY category description
To be able to dress dolls	d155	Acquiring skills
	d131	Learning through actions with objects
	d4402	Manipulating
	d880	Engagement in play
To be able to swim more efficiently, with smoother and coordinated action	d4554	Swimming
To be able to toilet independently	d530	Toileting
To be able to bath and shower independently	d5101	Washing whole body
To be able to carry a tray (in a café, canteen, to transport meals)	d4301	Carrying in the hands
To be able to fasten buttons and zips	d4402	Manipulating
To be able to gel hair	d5202	Caring for hair

To be able to operate and use a mobile phone	d3600	Using telecommunication devices
	d4402	Manipulating
To be able to operate buttons on a remote control device e.g. TV or DVD player	d4402	Manipulating
To be able to play ball games (throw a ball)	d4454	Throwing
	d9201	Sports
	d4452	Reaching
To be able to play board games	d131	Learning through actions with objects
	d155	Acquiring skills
	d2103	Undertaking a single task in a group
	d4403	Releasing
	d4452	Reaching
	d9200	Play
To be able to wash and bath independently	d5101	Washing whole body
To be able to use stationery	d1450	Acquiring skills to use writing implements

To close and open doors using affected arm	d4450	Pulling
	d4451	Pushing
To improve spontaneous use of limb/reduce neglect	b1801	Body image
To improve use of a computer keyboard	d1450	Acquiring skills to use writing implements

Spasticity Clinic audit goals and outcomes

Goal	**ICF-CY code**	**ICF-CY category description**
Help with dressing	d540	Dressing
Improve ability to grasp	d4401	Grasping
Improve arm function to help sign more effectively	d340	Producing messages in formal sign language
Improved pincer grip	d4402	Manipulating
Improved voluntary grip	d4401	Grasping
Improve cosmetic appearance (of arm)	b180	Experience of self and time functions
Improve joint ROM	b710	Mobility of joint functions
Improved finger movement	b7101	Mobility of several joints

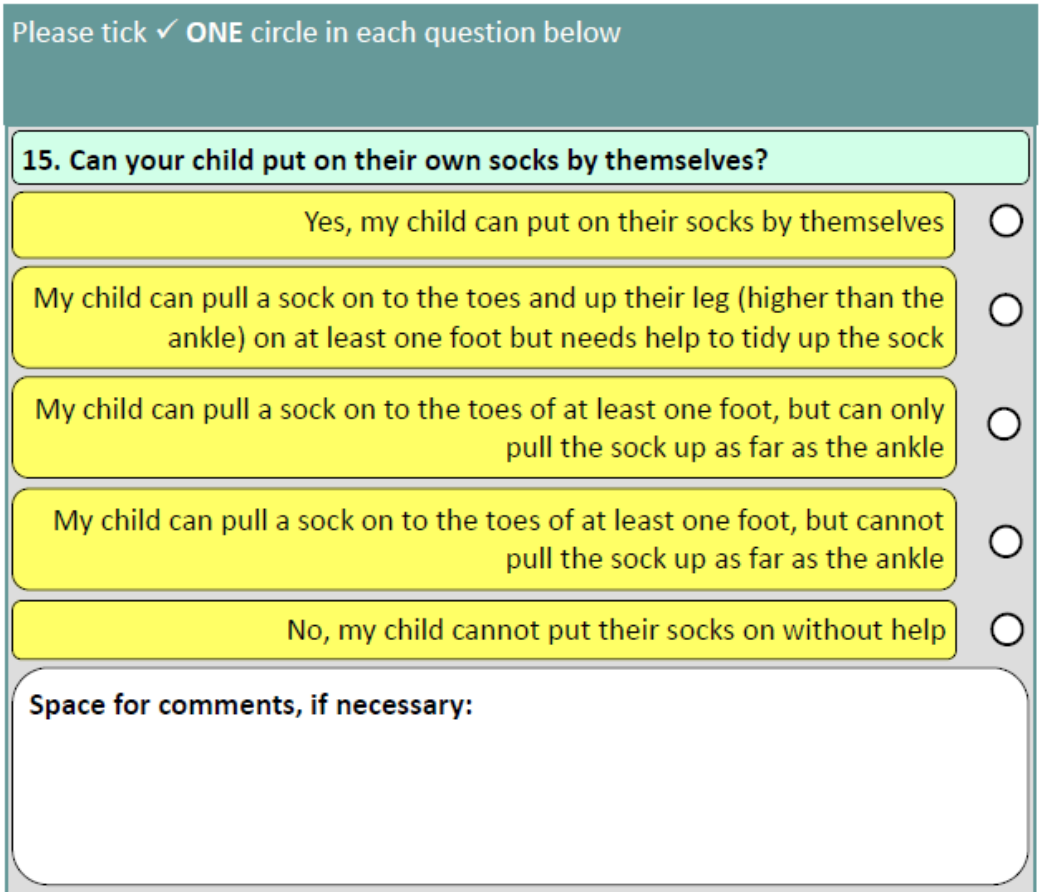
Using arm in more functional ways	b760	Control of voluntary movement functions
Able to hold reins when horse-riding	d440	Fine hand use
Assist in ADL	d630-d649	Household tasks
Improve functional reach	d4452	Reaching
Improve hand function	d440	Fine hand use
Able to put shoes on more easily	d2104	Completing a simple task
	d5402	Putting on footwear
Easier to crawl	d4550	Crawling
Improve hygiene	d510	Washing oneself
Improve supination	b710	Mobility of joint functions
Reduce flexed elbow posturing	b7100	Mobility of a single joint

Abbreviations

ICF-CY: International Classification for Functioning, Disability and Health for Children and Youth;

An illustration of an item and its response categories from the ChARM questionnaire is shown in Figure 3-6.

Figure 3-6. Item 15 from the ChARM questionnaire before psychometric testing.



The final 40 items included in the ChARM questionnaire are given in Table 3-12.

Table 3-12. The final set of 40 items, mapped to ICF-CY categories.

Item number	Item	ICF-CY code	ICF-CY category
Item 1	Can your child reach out to touch you with both hands when facing you?	d4452	Reaching
Item 2	Can your child turn on a room light using a light switch on the wall, even if they have to use something (anything) to help them?	d4452 d2104 d4402	Reaching Completing a simple task Manipulating
Item 3	Can your child make purposeful hand gestures? For example, would your child be able to communicate using hand signs (like Makaton sign language) or gestures?	d335 d350	Producing nonverbal messages Conversation
Item 4	Does your child ignore their less-preferred arm?	b1801	Body image
Item 5	Does your child use their less-preferred arm naturally?	b1801	Body image
Item 6	Can your child turn the palm of their preferred hand upwards to receive a treat (e.g. a sweet) into their palm?	d4452 d440	Reaching Fine hand use
Item 7	Could your child pick up a coin from a table with one hand and put it into a purse or wallet held using the other arm or hand?	d1201 d131	Touching Learning through actions with objects

		d2102	Undertaking a single task independently
		d4402	Manipulating
		d4403	Releasing
Item 8	Can your child button a polo shirt (one that only has a few buttons)?	d4402	Manipulating
		d540	Dressing
Item 9	Can your child move pieces around a games board e.g. Snakes and Ladders, Draughts, Trivial Pursuit, Monopoly, Solitaire or other board games?	d131	Learning through actions with objects
		d155	Acquiring skills
		d2103	Undertaking a single task in a group
		d4403	Releasing
		d4452	Reaching
		d9200	Play
Item 10	Can your child use a computer keyboard?	d1450	Acquiring skills to use writing implements
Item 11	Can your child write their name using a pen, pencil or crayon?	d170	Writing
		d1450	Acquiring skills to use writing implements
		d440	Fine hand use
Item 12	Can your child clean their own teeth, using any kind of toothbrush, if the toothpaste is put on the brush for them?	d1550	Acquiring basic skills
		d4453	Turning or twisting the hands or arms

		d5201	Caring for teeth
Item 13	Is your child able to open a previously opened jar of their favourite spread e.g. chocolate spread, peanut butter or jam?	d4402	Manipulating
		d4401	Grasping
		d4403	Releasing
		d4453	Turning or twisting the hands or arms
Item 14	Can your child feed themselves using a spoon?	d4402	Manipulating
		d550	Eating
		d1550	Acquiring basic skills
Item 15	Can your child put on their own socks by themselves?	d2104	Completing a simple task
		d5402	Putting on footwear
Item 16	Can your child completely wash his or herself in the bath or shower?	d5101	Washing whole body
Item 17	Can your child use your mobile phone?	d3600	Using telecommunication devices
		d4402	Manipulating
Item 18	Can your child move independently from their bed to a chair?	d410	Changing basic body position
		d4201	Transferring oneself while lying
Item 19	Can your child pour breakfast cereal into a bowl from a box of cereal that is already open (e.g. Cheerios, Frosties, Cornflakes)?	d4401	Grasping
		d4403	Releasing

		d550	Eating
		d6300	Preparing simple meals
Item 20	Can your child spread butter (or margarine) on a slice of bread?	d4402	Manipulating
		d6300	Preparing simple meals
Item 21	Can your child use a remote control to operate the TV or DVD player?	d4402	Manipulating
Item 22	Can your child zip up a coat by themselves?	d4402	Manipulating
Item 23	Can your child put on a pair of long trousers if they are laid out properly for them first?	d4403	Releasing
		d4401	Grasping
		d540	Dressing
Item 24	Can your child open a car door?	d4402	Manipulating
		d4453	Turning or twisting the hands or arms
Item 25	Can your child wash and dry their hands?	d510	Washing oneself
		d5100	Washing body parts
		d4402	Manipulating
Item 26	Can your child throw a tennis ball (or similar-sized ball) to a catcher?	d4454	Throwing
		d9201	Sports
		d4452	Reaching

Item 27	Can your child catch something thrown from three steps away?	d4454	Throwing
		d4403	Releasing
		d9201	Sports
Item 28	Can your child put on a vest (or short-sleeved T-shirt - don't worry about buttons) if it is laid out properly for them?	d4402	Manipulating
		d540	Dressing
Item 29	Can your child tidy their bedroom?	d630-d649	Household tasks
Item 30	Can your child get into and out of the bath without help?	d410	Changing basic body position
		d420	Transferring oneself
Item 31	Can your child pour a drink e.g. milk or juice from an opened carton into a beaker or cup?	d4401	Grasping
		d4403	Releasing
		d560	Drinking
Item 32	Can your child go the toilet by themselves (that is, undress, use the toilet, clean themselves, dress themselves and wash their hands)?	d4403	Releasing
		d4401	Grasping
		d540	Dressing
		d4402	Manipulating
		d530	Toileting
Item 33	Can your child pick up a coin from a table and put it into a money box?	d1201	Touching

		d131	Learning through actions with objects
		d2102	Undertaking a single task independently
		d4402	Manipulating
		d4403	Releasing
Item 34	Can your child use a ruler for drawing and for underlining words?	d1450	Acquiring skills to use writing implements
Item 35	Can your child complete their homework using a pen or pencil (not a computer)?	d1450	Acquiring skills to use writing implements
		d440	Fine hand use
Item 36	Can your child swim?	d4554	Swimming
		d9201	Sports
Item 37	Can your child pick up and hold a plate or tray of food?	d4301	Carrying in the hands
Item 38	Can your child use both hands when writing or drawing e.g. one hand to write or draw and the other to hold the book open or the paper still?	d1450	Acquiring skills to use writing implements
Item 39	Can your child crawl on hands and knees independently?	d4550	Crawling
Item 40	Can your child apply hair products to their hair independently (e.g. shampoo or hair gel)?	d5202	Caring for hair

Abbreviations: ICF-CY: International Classification for Functioning, Disability and Health for Children and Youth;

3.4.4 Discussion of development of the ChARM questionnaire

There are a number of psychometric and practical limitations among the measures commonly used to evaluate upper limb activity limitation in children with cerebral palsy. Testing measures with modern psychometric techniques can provide guidance to correct these limitations during (re)development of measures but nothing can rectify poorly designed or chosen content, such as questionnaire items. Rather than adopt or adapt items from current measures new items were developed from goals of activity limitation rehabilitation developed by clinicians and researchers for upper limb rehabilitation programmes involving children with cerebral palsy. This novel approach was expected to obtain the most common activities at which children with cerebral palsy experience activity limitations, and to establish confidently that the items for the new measure were validated for children with cerebral palsy aged between five and sixteen years. The large number of duplicate goals suggests that the approach met with some success.

Given the evidence that reliability and responsiveness increase with the number of response categories, more than three response categories were intended for each item. It was also decided not to adopt a standard (identical) response format for each item. Instead each item's activity was broken down into natural stages of achievement which would be easily identifiable to the respondent completing the questionnaire. The number of response options to each item was therefore dependent on the number of natural stages of achievement of the activity and the number of response options varied for each item. All potentially appropriate response options were included, based on confidence in the likelihood that the future Rasch analysis will identify disordering of thresholds and show which responses are working and which are not. Even so, this process produced some items with three or fewer natural response categories.

Although this approach means that the ChARM will be more time-consuming to complete for respondents (because each item has different responses to read and consider), it offers several advantages. Firstly, the optimal number of response options has been generated for each item. Secondly, it facilitates easy identification of stages of achievement which a child has reached. It also avoids the potential uncertainty for the respondent of which response option to endorse that occurs with homogenous item response options, and

finally it potentially solves the problem of the halo effect (when respondents endorse the same response category for each item).

A number of goals were returned that were not within the ICF-CY activity and participation domain but fell within the domain of body functions and structures. Examples of these included “spontaneous use of the affected or more affected arm”, “to reduce flexed elbow posturing”, or “to improve supination”. These were developed into items because they were goals of upper limb activity rehabilitation, developed to reduce the impairment underlying the activity limitation, and could be broken down into natural stages of achievement to form responses to an item stem. Their appropriateness for inclusion in the final draft of the ChARM will be determined by the psychometric testing, which includes identification of ill-fitting items.

Recent studies investigating the efficacy of new approaches to reducing activity limitation (e.g. Sakzewski et al., 2011a, Wallen et al., 2007) have independently identified the same or similar goals used in the development of the ChARM, suggesting that my efforts to identify the most common activity limitations in children with cerebral palsy have met with some success. However, it is recognised that a smaller range and breadth of goals was received than had been expected, and the final sample size of 53 children was substantially smaller than anticipated, given that the therapy teams involved covered a well-populated area potentially including up to 2,000 children with cerebral palsy, none of whom were excluded outside of the age range five to sixteen years. Possible reasons for this poor initial recruitment include the requirement of all participants to give full, written, informed consent to participating in the study despite the low impact of the study on children's care. This has now been recognised by the NHS research ethics service and proportionate review is now available for studies of this nature. Therapy teams were also impacted by financial cutbacks in their services which affected their ability to participate in research studies: for example, one therapy team that offered to support the study was only offering phone and written advice to patients, and were not having face-to-face contact due to staffing shortages. Additionally, an unknown number of parents were excluded from participation based on the judgement of participating paediatric teams.

However, although this number of children is smaller than anticipated, 78 unique goals delivered a wide range of appropriate ICF-CY activity-related categories. The final item set has a broad range of activities and includes items which are potentially achievable by some of the most disabled children with cerebral palsy (e.g. Items 1 and 3, see Table 3-12 above), thus reducing any potential floor effect.

3.4.5 Conclusion of development of the ChARM questionnaire

A critical appraisal and an overview of psychometric techniques identified a number of problems with measures of upper limb activity limitation. The ChARM was conceptualised through the realisation that careful and methodological development of items and their response categories, followed by psychometric testing using Rasch analysis, could overcome these limitations. The items were developed using a novel and comprehensive method, and ensured validity and appropriateness for the population and health condition for which use of the measure is intended. The outcome is a questionnaire of 40 items with established face validity that represent a range of ICF-CY activities with which children with cerebral palsy commonly experience limited achievement. The questionnaire now requires psychometric testing.

3.5 Psychometric testing of the ChARM questionnaire – Rasch analysis

Part 2 of ChARM development uses Rasch analysis to evaluate the psychometric properties of the ChARM questionnaire. The findings of this analysis e.g. disordered response categories and items which do not fit the Rasch model, will guide amendments to the questionnaire for the purpose of producing a psychometrically valid measure. While the development of the ChARM was designed and carried out in a way to maximise the likelihood of item validity and unidimensionality i.e. a measure of only upper limb function in children with cerebral palsy, the Rasch analysis will identify multidimensionality of items.

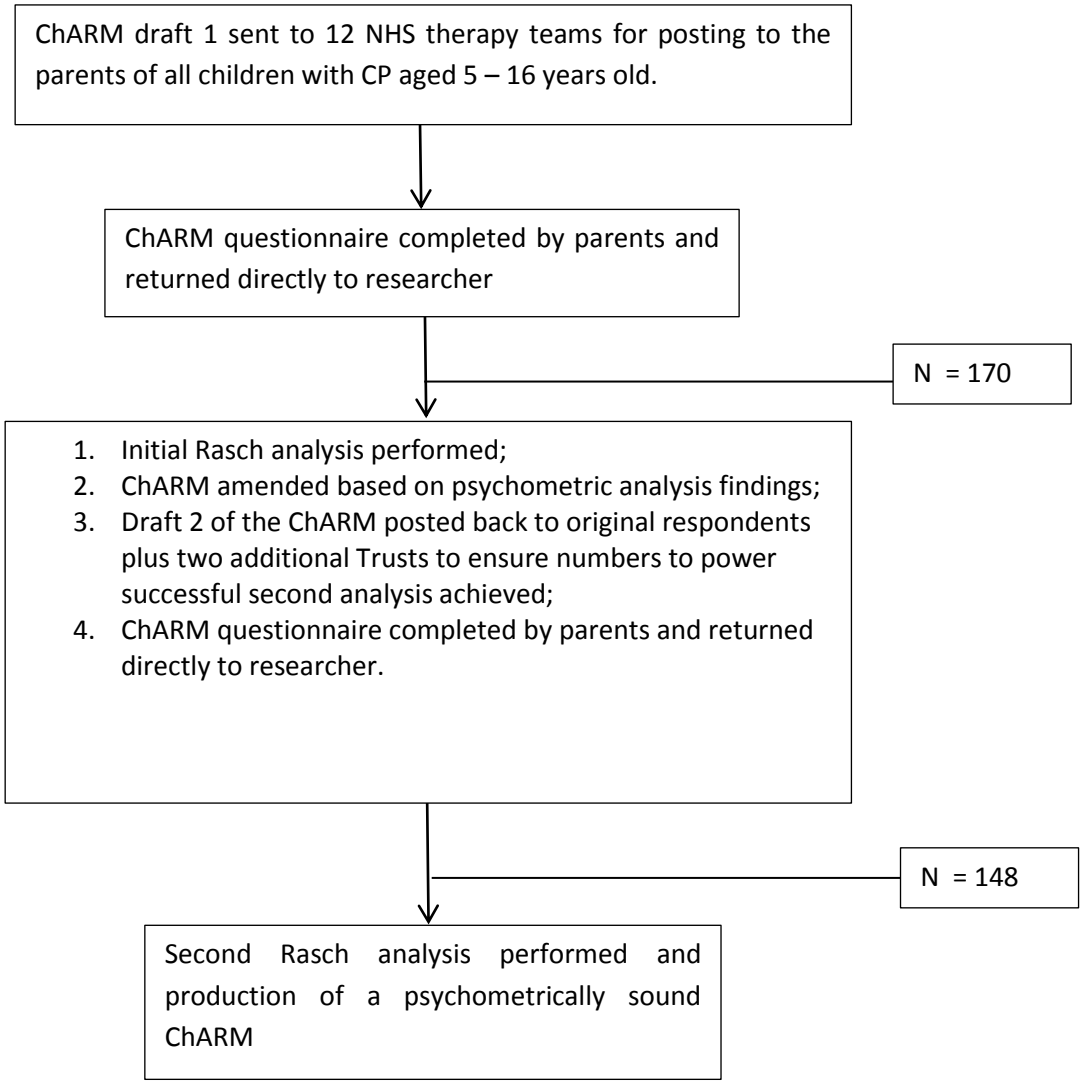
This section describes the Rasch analysis of the ChARM questionnaire and its final development and preparation for publication. This procedure takes place in two stages: firstly, the 40-item questionnaire undergoes Rasch analysis and modification after completion by 170 parents of children with cerebral palsy aged 5 - 16 years old; then the resulting questionnaire will be completed by 148 parents of children with cerebral palsy aged 5 - 16 years old and undergo a final Rasch analysis and modifications. The aim is to produce a measure of upper limb activity limitation that is validated for children with

cerebral palsy aged 5 - 16 years old, and that is unidimensional, responsive to changes in activity limitation, will overcome the problem of ceiling or floor effects and will allow the transformation of the ordinal raw scores into linear outcome scores.

3.5.1 Method

The conduct of the study is shown in Figure 3-7.

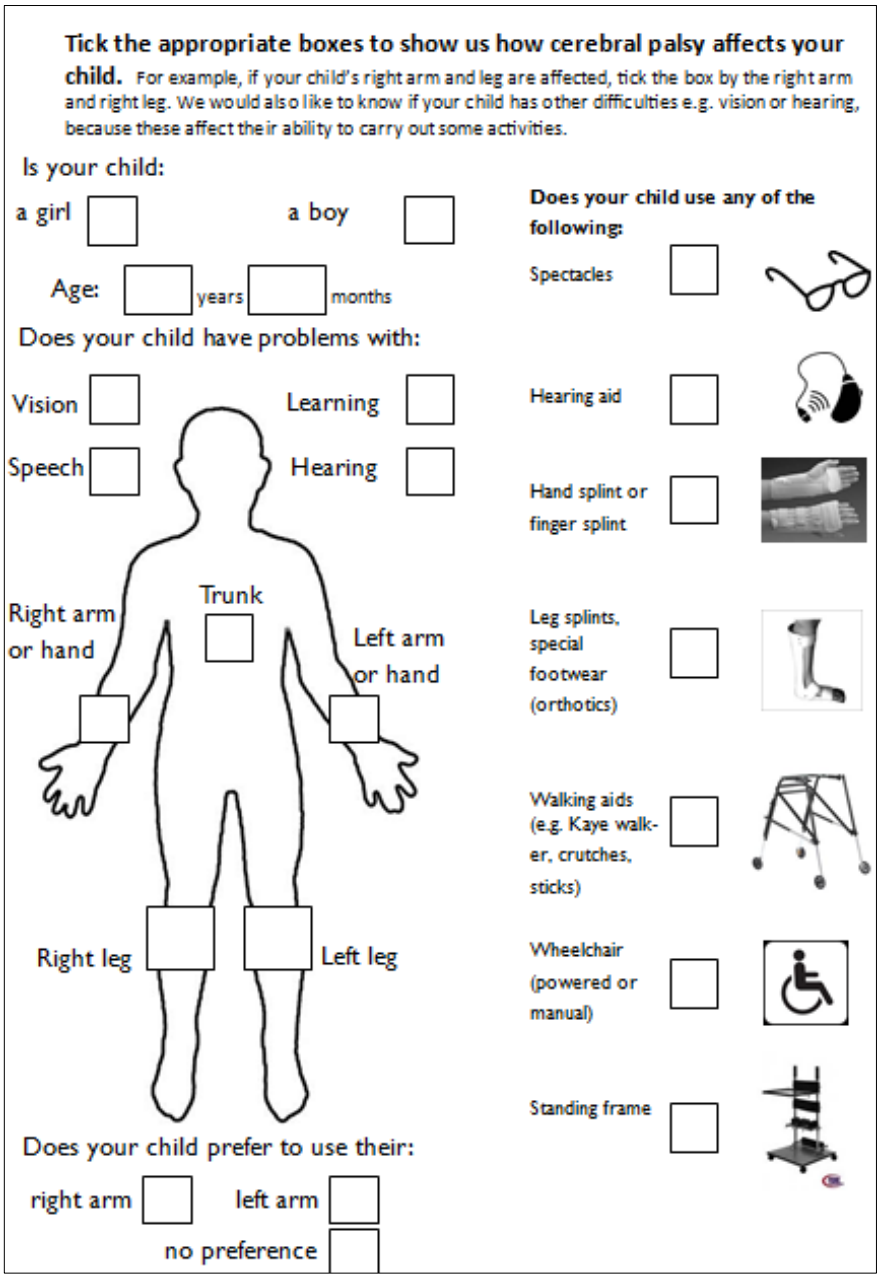
Figure 3-1. Conduct of sub-study to perform psychometric testing (Rasch analysis) of ChARM.



Therapy teams from the nine NHS Trusts in Table 3-6 on page 147 and the three NHS Trusts in Table 3-9 on page 153 posted the first draft of the 40-item ChARM to parents of all children with cerebral palsy aged 5 – 16 years old that were known to the therapy teams, along with a REC-approved information sheet and a postage-paid pre-addressed envelope for the completed ChARM to be returned directly to the Chief Investigator. The

ChARM included the form shown in Figure 3-8 for parents to return clinical and demographic information that is essential for comprehensive psychometric testing e.g. to assess Differential Item Functioning (DIF). A tick box form was also included for parents to indicate the manual ability of their child using the MACS. Social media (Facebook and online forums on Scope and HemiHelp websites) were used to advertise an online version of the questionnaire which could be either downloaded and posted to the researchers or completed electronically. In order to power an adequate Rasch analysis a minimum of 150 completed ChARMs were required to achieve 99% confidence of item calibration to within 0.5 logits (Linacre, 1994).

Figure 3-8. Mannequin for parents to give clinical details of their child.



3.5.2 Initial Rasch analysis of ChARM draft 1 (development of draft 2)

Completed ChARMs were received from 170 parents of children with cerebral palsy, achieving the minimum of 150 datasets required for an adequate Rasch analysis. Rasch analysis was carried out using RUMM2030 Version 5.4 for Windows, Copyright 1997 – 2012 Rumm Laboratory Pty Ltd. The Masters Partial Credit Model (Unrestricted; Polytomous/Extended Response Category test format) (Masters, 1982) was used because item responses varied in type and number between items, and differences between item thresholds were anticipated (Tennant and Conaghan, 2007).

Summary statistics for the initial analysis are given in Table 3-13 below. Fit of items to the Rasch model were evaluated using χ^2-statistics, where a statistically significant result meant that the items were significantly different from the model prediction (that is, an acceptable fit to the Rasch model was shown by a non-significant p-value). The 40 items of draft 1 of the ChARM therefore presented a poor fit to the model, with a χ^2-statistic of 647 and a p-value of less than 0.001 (df = 80).

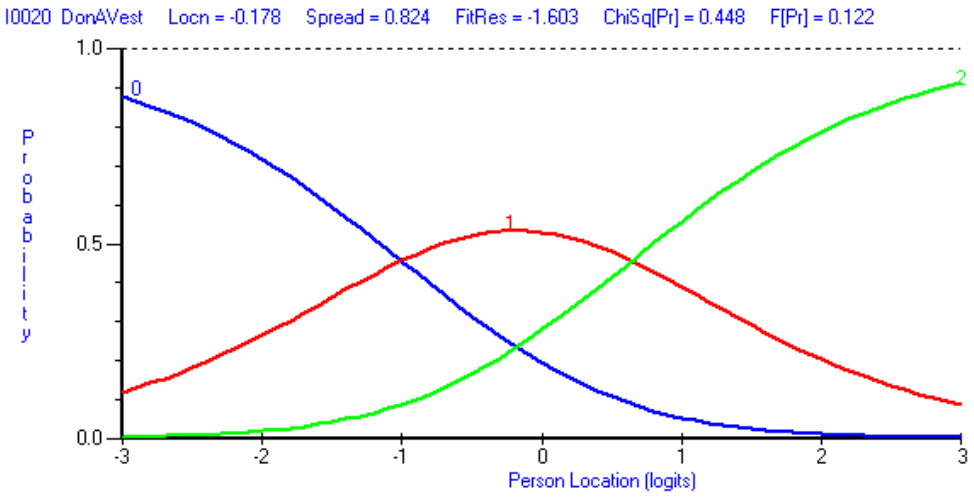
Table 3-13. Initial Rasch analysis of the ChARM, first draft.

	Item Location		**Person Location**		**Item Fit Residual**		**Person Fit Residual**		**Chi Square Interaction**			**Reliability**			**Unidimensionality t-tests (CI)**			
												PSI			number of			lower bound
Analysis	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Value	df	p	with extrms	No extrms	Alpha	significant tests	out of:	%	95% CI
Initial analysis of draft 1 received from 170 parents	0.00	0.80	0.10	1.77	-0.36	2.81	-0.12	1.08	647	80	<0.001	0.97	0.97	0.95	36	169	21	0.18

3.5.2.1 Disordered response categories

The first psychometric problem to be addressed was the disordering of response category thresholds. Figure 3-9 below shows response categories thresholds that are satisfactorily ordered. Ability is shown on the x-axis, and the response categories are shown as response curves, one curve for each category. As the level of ability increases, each response categories is endorsed in turn, because each category indicates an increasing level of ability (and in the case of the ChARM, a decreasing level of activity limitation). Note that the thresholds, where the response category curves intersect, are properly ordered in increasing levels of ability.

Figure 3-9. An example of ordered response categories.



In the case of Figure 3-9, the first response category threshold is at -1 logit, and the second response category threshold is at 0.7 logits. A child who was below -1 logit would have the first response category endorsed; children between -1 logit and 0.7 logit would have the second category endorsed; and above 0.7 logits children would have the third category endorsed.

Figure 3-10. Disordered response categories.

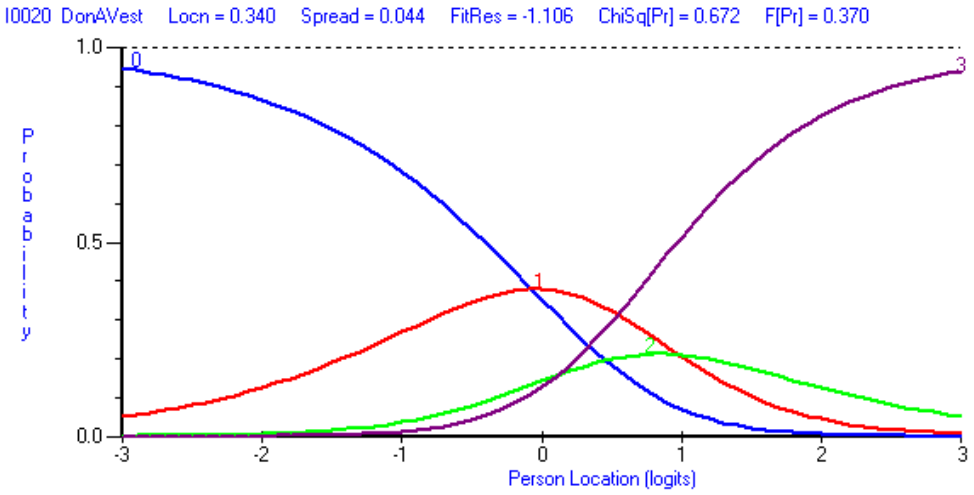


Figure 3-10 shows disordered response category thresholds. In this case, the threshold for the intersection of response category curves 1 and 2 meet at 1 logit, but the response category threshold for curve 2 intersects with response category curve 3 at 0.1 logit. The disordering of response categories occurs when categories are unclear, or there are too many responses so that respondents (in this case, the parents) endorse adjacent responses even though the child with cerebral palsy has similar levels of activity limitation. Because of the ordinal nature of the response categories, this could mean that one parent is indicating that their child has a lower degree of activity limitation than a second parent, even though the second parent's child in fact has a greater degree of limitation.

Addressing disordered thresholds involves collapsing or combining response categories. This is explained in more detail later in this section but in the case illustrated above, combining the response categories for curves 1 and 2 in Figure 3-10 produced the response curves in Figure 3-9 on page 181.

Fifteen items showed disordered thresholds after the first analysis of the ChARM draft 1, which were corrected by collapsing categories as described.

3.5.2.2 Item fit

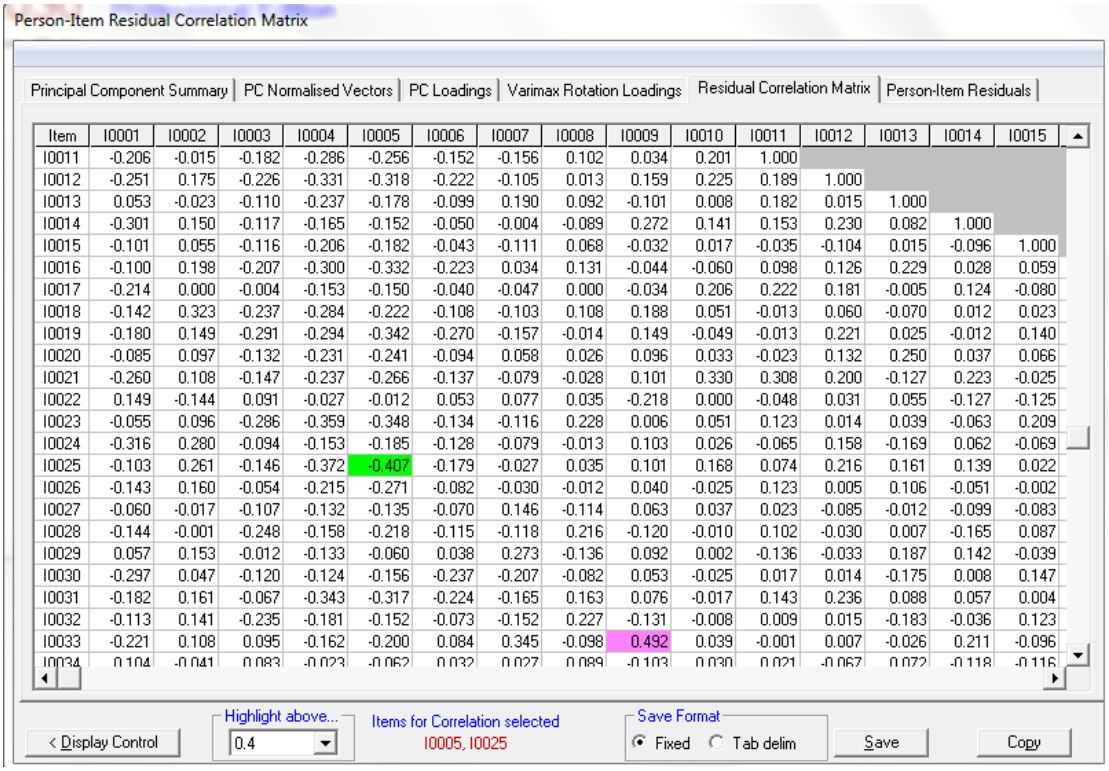
Individual item fit statistics revealed that a number of items did not fit the Rasch model. Item 3, item 4 and item 5 showed a high χ^2-statistic with significant p-values. They also showed an unacceptably high *positive* residual value – this indicates that the items are

unrelated to the variable being measured. Items 5 and 6 were included because they were notable from our earlier research as items that indicated the child was overcoming learned non-use, but it was not unexpected to see that they did not fall within acceptable statistical parameters. Item 4, item 16, item 23, item 25 and item 31 are similar but have high *negative* residuals, indicating that they redundant, with a high likelihood of being influenced strongly by other items. These seven items were selected for deletion.

3.5.2.3 Local dependency

Figure 3-11 below shows part of the correlation matrix produced by the RUMM software program, in which item 5 is correlated with item 25, and item 9 and item 33 are correlated.

Figure 3-11. A snapshot showing part of the correlation matrix that suggests local dependency of items.



Correlation of items suggests local dependency. Table 3-14 below lists all items that were highlighted in the correlation matrix and shows which items were selected for deletion. When considering items for deletion, the emphasis was on retaining items that describe bilateral manual activity and which were considered more inclusive of children with a greater disability (MACS level IV and V).

Table 3-14. Items showing local dependency.

Correlated items	Item to be deleted	Comments
Item 4 and item 5	Both items	Items deleted because of poor fit to the Rasch model
Item 23 and item 28	Item 23	Item deleted because of poor fit to the Rasch model
Item 9 and item 33	Item 33	
Item 7 and item 33	Item 33	
Item 10 and item 21	Item 21	
Item 17 and item 21	Item 21	
Item 23 and item 32	Item 23	Item deleted because of poor fit to the Rasch model
Item 19 and item 31	Item 31	Item deleted because of poor fit to the Rasch model
Item 18 and item 32	Item 32	
Item 16 and item 25	Item 25	Item deleted because of poor fit to the Rasch model
Item 12 and item 35	Item 35	
Item 11 and item 35	Item 35	

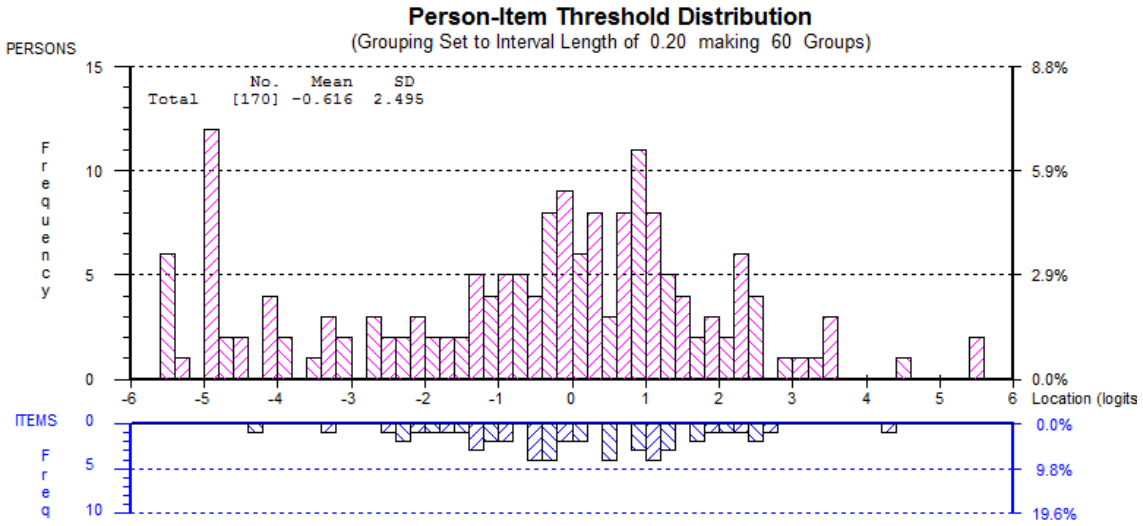
3.5.2.4 Unidimensionality

Unidimensionality was tested by finding the two most divergent subsets of items derived from the first residual component (factor) using a principal components analysis as described in Tennant and Conaghan (2007). Person estimates were calculated from two subsets that were made up of the items identified by those that loaded most positively (one subset) and negatively (the second subset) on the first factor. Differences between them were determined using a series of t-tests. Then, a binomial test of proportions showed that the number of t-tests (7 out of 149, or 4.7%) significant at 5% overlapped the 5% lower confidence interval, suggesting that the ChARM draft 2 was unidimensional.

3.5.2.5 Final psychometric testing and ChARM draft 2 production

When the items that had so far been identified as causing psychometric problems were deleted, a Rasch analysis was repeated. Eight items showed misfit and were deleted. Following this procedure, the measure consisted of 19 items showing acceptable fit to the Rasch. However, twenty-two children fell outside of the person-item thresholds, as illustrated in Figure 3-12 (three at the upper end of the threshold distribution and nineteen at the lowest end). This represents a floor effect.

Figure 3-12. Person-item threshold distribution for the ChARM after first Rasch analysis and development.



Final summary statistics for the ChARM draft 2 are shown in Table 3-15 below. The ChARM draft 2 consists of 19 items which show acceptable fit to the Rasch model (χ^2-statistic of 46 and a p-value of less than 0.19, df = 38), are unidimensional and have no local dependency, but show a large floor effect.

Table 3-15. Final summary statistics for the second draft of the ChARM.

	Item Location		Person Location		Item Fit Residual		Person Fit Residual		Chi Square Interaction			Reliability			Unidimensionality t-tests (CI)			
												PSI						
Analysis	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Value**	**df**	**p**	with extrms	NO extrms	**Alpha**	**number of significant tests**	**out of:**	**%**	**lower bound 95% CI**
Final analysis of draft 1 after amendment of questionnaire items	0.00	1.10	-0.62	2.50	-0.01	1.12	-0.16	0.93	46	38	0.19	0.94	0.94	0.95	7	149	4.7	0.012

One of the 170 parents returning the ChARM questionnaire for the initial Rasch analysis was critical of the items, suggesting that they were all too difficult for children with activity limitations of the severity experienced by her child. The parent's child, whom she rated as MACS Level V, had fallen within the floor effect exhibited by the ChARM. This parent included within the returned ChARM questionnaire a number of activities that the parent wanted as rehabilitation goals for her daughter. With the agreement of the parent, the activities were written into six items with response categories. The parent responded favourably about the new items and suggested that her child would be able to partly achieve some of them.

The six items were included in the ChARM draft 2, which now consisted of the 25 items shown in Table 3-16. The final stage was therefore a repeat of the psychometric analysis, but on the ChARM draft 2 questionnaire once it was completed by a sufficient number of parents of children with cerebral palsy.

Table 3-16. Items included in the second draft of the ChARM, and the number of response categories for each item.

Item number	Item	Number of response categories
Item 1*	Can your child use their arm or hand to rub their nose e.g. when they have an itch or want to wipe it?	5
Item 2*	Can your child use their hand to press a switch or button e.g. to switch on a light or toy, or to make something happen?	3
Item 3*	Can your child clasp their hands together in front of them?	4
Item 4*	Can your child use their arms and hands to push things away in play e.g. a toy with wheels, or bubbles, a ball or balloon?	4
Item 5*	Can your child use their arms or hands to push away an activity or object that they don't like?	3
Item 6*	Can your child gather in clothes, towels, blankets or a soft toy with their arms and hands to clasp to their chest, either to hold for comfort or to carry?	4
Item 7	Can your child pick up a one pound coin from a table with one hand and put it into a purse or wallet held in the other arm or hand?	4
Item 8	Can your child button a polo shirt (one that only has a few buttons)?	3
Item 9	Can your child move pieces around a games board e.g. Snakes and Ladders, Draughts, Trivial Pursuit, Monopoly,	4

	Solitaire or other board games?	
Item 10	Can your child use a computer keyboard?	3
Item 11	Can your child clean their own teeth, using any kind of toothbrush, if the toothpaste is put on the brush for them?	3
Item 12	Can your child open a previously opened jar of spread e.g. chocolate spread, peanut butter or jam?	3
Item 13	Can your child feed themselves using a spoon?	4
Item 14	Can your child put on their own socks by themselves?	4
Item 15	Can your child pour breakfast cereal into a bowl from a box of cereal that is already open (e.g. Cheerios, Frosties, Cornflakes)?	4
Item 16	Can your child spread butter (or margarine) on a slice of bread?	5
Item 17	Can your child zip up a coat by themselves?	3
Item 18	Can your child throw a tennis ball (or similar-sized ball) to a catcher?	2
Item 19	Can your child catch something thrown from 3 steps away?	4
Item 20	Can your child put on a vest (or short-sleeved T-shirt - don't worry about buttons) if it is laid out properly for them?	4
Item 21	Can your child use a ruler for drawing and for underlining words?	4
Item 22	Can your child tidy their bedroom?	4
Item 23	Can your child pick up and hold a plate or tray of food?	3
Item 24	Can your child use both hands when writing or drawing e.g. one hand to write or draw and the other to hold the book open or the paper still?	4
Item 25	Can your child apply hair products to their hair independently (e.g. shampoo or hair gel)?	3

* indicates the items added to the second draft that were suggested by a parent who completed draft 1 as appropriate for her child who was Manual Ability Classification System Level V.

3.5.4 Rasch analysis of draft 2 and production of ChARM Version 1

The modified questionnaire, ChARM draft 2, was posted back to all 170 parents that had returned the first draft. A postage-paid and pre-addressed envelope was included. To overcome the possibility that not every parent would return a questionnaire, therapy teams from two additional Trusts (see Table 3-17) posted out the questionnaire to parents of children with cerebral palsy aged 5 – 16 years known to the teams.

Table 3-17. Two additional paediatric therapy teams participating by posting out the ChARM to parents of children with cerebral palsy.

PARTICIPATING NHS TRUST THERAPY TEAM
Humber NHS Foundation Trust
Chesterfield Royal Hospital NHS Foundation Trust

3.5.4.1 Results

For the final psychometric testing of the ChARM 148 completed draft 2 questionnaires from parents of children with cerebral palsy were received. The children's demographics and clinical details are given in Table 3-18. Twelve of the children were reported by their parents as having minimal upper limb involvement (MACS Level I). All children were included in the psychometric testing.

Table 3-18. Demographics of the sample of children with cerebral palsy on whom the final psychometric testing is based.

Demographics (n=148)		
Age in years (y) and months (m)	mean (SD)	10y1m (3y3m)
	median	11y9m
	min	4y8m
	max	16y11m
Gender	Male	85 (57%)
	Female	57 (39%)
	missing data	(6) (4%)
MACS Levels	Level I	9 (6%)
	Level II	26 (18%)
	Level III	48 (32%)
	Level IV	45 (30%)
	Level V	18 (12%)
	missing data	(2) (2%)
Distribution	Bilateral	77 (52%)
	Unilateral	59 (40%)
	Lower limb only*	12 (8%)
Learning impairment	Present	85 (57%)
	Not present	61 (41%)
	missing data	(2) (2%)
Visual impairment	Present	62 (57%)
	Not present	84 (39%)
	missing data	(2) (2%)
Hearing impairment	Present	18 (12%)
	Not present	128 (86%)
	missing data	(2) (2%)
Speech impairment	Present	72 (48%)
	Not present	74 (50%)
	missing data	(2) (2%)

MACS: Manual Ability Classification System; SD: standard deviation;

y = years; m = months.

*12 parents reported children as lower limb impairment with no upper limb involvement; these children were included within the analyses

Initial summary statistics for the draft 2 ChARM are given in Table 3-19 below. These figures suggest that the ChARM shows a large misfit to the Rasch model (χ^2-statistic = 129, df 50, p <0.001).

Table 3-19. Summary of statistics for the initial Rasch analysis of the ChARM draft 2.

Item Location		**Person Location**		**Item Fit Residual**		**Person Fit Residual**		**Chi Square Interaction**		
Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Value**	**df**	**p**
0.00	1.55	1.00	2.35	-0.41	0.95	-0.25	1.06	129	50	<0.001

Reliability			**Unidimensionality t-tests (CI)**			
PSI						
with extrms	NO extrms	**Alpha**	**number of significant tests**	**out of:**	**%**	**lower bound 95% CI**
0.96	0.96	0.95	29	146	20	0.163

3.5.4.1.1 Disordered thresholds

Items 1, 6, 13, 14 and 20 showed disordered response thresholds. Disordered thresholds were corrected by collapsing categories, as shown for items 6 and 20 in Table 3-20. The disordered thresholds for item 20 are further illustrated by the category response curves in Figure 3-10 on page 182. Combining the responses as shown in Table 3-20 produced the ordered thresholds shown in Figure 3-9 on page 181. Disordered responses for item 1, item 6, item 13 and tem 14 were also collapsed in a similar way. Items 1 and 14 were subsequently deleted entirely to establish acceptable unidimensionality and fit to the Rasch model).

Table 3-20. Collapsing response categories to correct disordered response thresholds.

Item 6	**Can your child gather in clothes, towels, blankets or a soft toy with their arms and hands to clasp to their chest, either to hold for comfort or to carry?**	
	Original item responses	Responses after combining item responses 2 and 3
Response 1	Yes, my child can easily gather things in to hold to their chest	Yes, my child can easily gather things in to hold to their chest
Response 2	My child can gather things in but only using one arm or hand	My child can gather things in to their chest but it is quite difficult
Response 3	My child can gather things in to their chest but it is quite difficult	
Response 4	No, my child cannot use their arms to gather things in to their chest	No, my child cannot use their arms to gather things in to their chest
Item 20	**Can your child put on a vest (or short-sleeved T-shirt - don't worry about buttons) if it is laid out properly for them?**	
	Original item responses	Responses after combining item responses 2 and 3
Response 1	Yes, my child can put on a vest (or T-shirt) by themselves	Yes, my child can put on a vest (or T-shirt) by themselves
Response 2	My child can correctly put on the vest (or T-shirt) but may need help to tidy it up at the back or arms	My child can almost put on a vest (or T-shirt) but needs help to tidy it up or finish putting it on
Response 3	My child can partly put on a vest (or T-shirt) but needs help to finish putting it on	
Response 4	No, my child needs full assistance to put on a vest (or T-shirt)	No, my child needs full assistance to put on a vest (or T-shirt)

3.5.4.1.2 Item fit

All items showed acceptable fit to the Rasch model except for item 2 (χ^2-statistic = 19.419, p = 0.000062 (Bonferroni-corrected)). All items showed acceptable residual values, meaning that there were no immediate indications of redundancy or that they were not contributing appropriately to the measure.

3.5.4.1.3 Unidimensionality

Using the procedure outlined above in 3.5.2.4 and described by Tennant and Conaghan (2007), a number of items were found to be contributing to a multidimensional scale. The

subset of items loading most positively onto the first factor were items 1 – 5, the negative subset were items 8, 13, 14, 20 and 22.

3.5.4.1.4 Local dependency

The correlation matrix indicated local dependency for a number of items, shown in Table 3-21.

Table 3-21. Draft 2 ChARM items showing local dependency.

Strongly correlated items	**Weakly correlated items**	
Item 4 and item 5	Item 3 and item 4	Item 3 and item 13
Item 5 and item 14	Item 2 and item 8	Item 8 and item 14
Item 14 and item 20	Item 4 and item 8	Item 5 and item 20
	Item 5 and item 8	Item 4 and item 22
		Item 5 and item 22

3.5.4.1.5 Production of ChARM version 1

The findings from the Rasch analysis given above were used to guide amendment of draft 2 of the questionnaire in order to produce version 1 of the ChARM. The emphasis was on retaining as many of the easiest items (items 1 – 6, item 18) as possible to attempt to overcome the floor effect. This was achieved though an iterative process that involved deleting items to achieve a fit to the model, then putting back easier items while at the same time avoiding local dependency pairings.

Table 3-22 below gives the final summary of statistics for the resulting version 1 of the ChARM, showing acceptable fit to the Rasch model (χ^2 (38) = 46, p = 0.18), acceptable unidimensionality and acceptable reliability. To achieve this, it was necessary to delete items 1 – 5 and item 14. The ChARM Version 1 is therefore a 19 item questionnaire.

Table 3-22. Summary of statistics for the final Rasch analysis of the ChARM version 1.

Item Location		Person Location		Item Fit Residual		Person Fit Residual		Chi Square Interaction		
Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Value**	**df**	**p**
0.00	1.40	-0.65	2.99	-0.18	0.98	-0.20	0.78	46	38	0.18

Reliability			Unidimensionality t-tests (CI)			
PSI						
with extrms	NO extrms	**Alpha**	**number of significant tests**	**out of:**	**%**	**lower bound 95% CI**
0.95	0.95	0.95	11	132	8	0.046

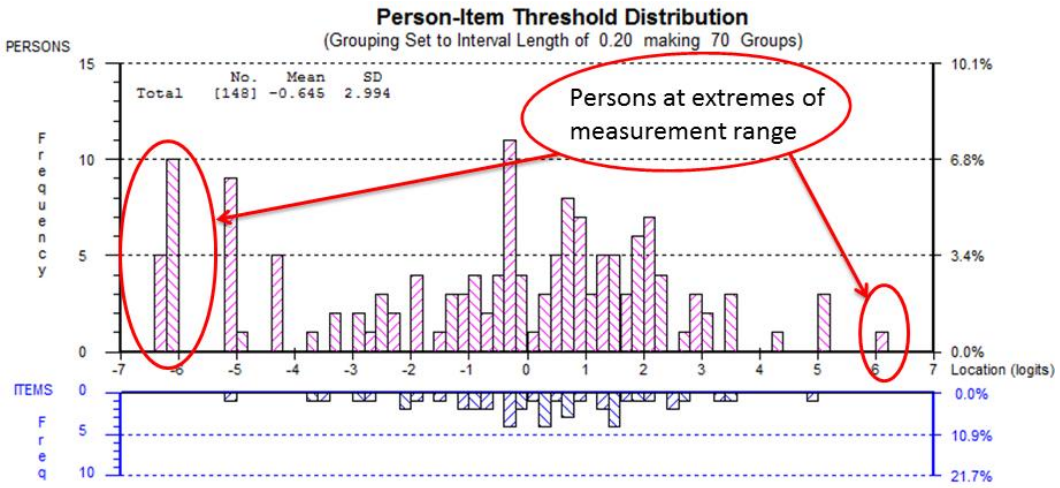
3.5.4.1.6 Final Rasch testing: differential item functioning (DIF)

The final version of the ChARM was tested for DIF in a number of groups: age, gender, distribution of cerebral palsy, visual impairment and learning difficulty. No DIF was present for any of these groups. This suggests, using gender as the example, that the items were appropriately and correctly used without bias based on whether the ChARM was completed for a male of female.

3.5.4.1.7 Floor and ceiling effects

The person-threshold distribution map in Figure 3-13 shows that 16 children fall outside of the measurement range of Version 1. This represents 12% of the sample, suggesting that there are no floor or ceiling effects in the ChARM.

Figure 3-13. Person-threshold distribution.



3.5.4.2 Discussion

Existing measures of upper limb activity limitation in children with cerebral palsy lack acceptable psychometric standards, possess characteristics such as difficult scoring or poor responsiveness, and can be expensive. The aim of this study was to develop a parent-completed questionnaire with the aim of producing a measure of upper limb activity for children with cerebral palsy aged five to sixteen years old that meets fundamental principles of scientific measurement. Additionally, it was aimed to overcome floor and ceiling effects, and to strive for maximum responsiveness to changes in levels of arm activity. Therefore, after a carefully considered, novel approach to item development, the Rasch model was applied to the first draft of 40 items and appropriate Rasch-guided modifications were carried out. The Rasch model was applied to the second draft, and modified again. The results suggest a successful outcome in developing a psychometrically sound measure with the properties that permit the transformation of its raw scores onto an interval level scale.

The main limitation with the psychometric testing of the ChARM is the low number of parents that completed each of the final two drafts of the ChARM, prior to the psychometric testing. However, a sample size was achieved that allows a strong initial calibration of items. All items in the ChARM fit the model and invariance is demonstrated by a scale-level non-significant χ^2 probability. Unidimensionality is acceptable. There is no apparent dependency between items at correlations of 0.3 and above, and no items show response bias (DIF) in clinical sub-groups of gender, age, distribution of limb impairment, learning or visual impairment.

Sixteen children (12%) are outside of the measurement range of the scale but this does not represent a floor effect or ceiling effect. Of the extreme scores, 15 were at the bottom of the scale. One of the extreme scores was from a questionnaire returned with no demographics or clinical data, 13 were MACS Level V, and one was MACS Level IV. Of the 14 with clinical data, all had learning disability and all but one had bilateral arm impairment. Five children of MACS Level V were represented on the scale, and these children also had learning disability. The extreme score at the top of the scale was a child with no learning disability and minimal unilateral upper limb impairment (11 years 11 months with MACS Level I). This child would therefore be expected to achieve a level of activity close to that of his able-bodied peers, and would not normally be referred for treatment by therapists. Therefore poor targeting at this demographic does not reduce the clinical utility of the ChARM.

The scale reliability (internal consistency) presented by the PSI and Cronbach's alpha is very high at 0.95. This value meets the standard required for use at the individual level and suggests that the ChARM will be responsive to small changes in arm activity.

Future testing of the ChARM should include test-retest reliability and evaluation of the responsiveness to change following intervention (in research and clinical practice). Future potential modifications to the ChARM include the addition of items to target the more disabled children with cerebral palsy. In addition, it is intended to develop an online scoring system that generates interval level outcome scores.

3.5.4.3 Conclusion

The ChARM is a measure of upper limb activity limitation validated for use with children with cerebral palsy aged five to sixteen years old. It is a parent-reported questionnaire of 19 items which describe activity limitation that is common to the majority of the population for which it is intended for use. Its ordinal outcome scores can be transformed to interval-level outcome scores. It has acceptable psychometric properties that suggest it will have adequate, clinically relevant responsiveness, but responsiveness and test-retest stability still require testing.

Following publication in a peer—reviewed journal, the ChARM will be made freely available for use by clinical staff and non-commercial researchers.

The ChARM was intended for use as the primary outcome measure for the RCT of the gaming device, but the process of development and testing took more than twice as long as expected. The production of the ChARM was therefore concluded in parallel with the RCT, and has not been used as an outcome measure in the RCT.

3.6 Assessment of upper limb kinematics of children with cerebral palsy: CPKAT

So far this chapter has highlighted the difficulties with the evaluation of activity limitation of children with cerebral palsy. These difficulties include a lack of responsiveness. Spatiotemporal analysis of upper limb movement has been suggested as a complementary assessment measure for the guidance and evaluation of cerebral palsy treatment because objective kinematic data provide fine-scale information that is not captured by the existing outcome measures discussed in sub-section 3.3 (Critical review of measures of upper limb

functional ability for children with cerebral palsy). Furthermore, it can improve planning for more appropriately-targeted treatments (Fitoussi et al., 2006, Fitoussi et al., 2011, Jaspers et al., 2011).

Sub-section 2.2 on page 50 (Design, construction and feasibility trial of prototype home-based assistive joystick and computer games (NIHR-funded study, grant ID G006)) includes the description and illustration of laboratory-based equipment for evaluating spatiotemporal movements of the upper limb of children with cerebral palsy. Because of the inconvenience for schoolchildren and their parents to undertake repeated travel to the laboratory for these evaluations it was necessary to explore other options for evaluating upper limb kinematics. Our solution was the adaptation of a tablet-based kinematic assessment tool called CKAT, which was developed and tested on non-disabled adults (Culmer et al., 2009). This adaptation is described and illustrated in sub-section 2.3.2.5.3 on page 70 and in Weightman et al. (2011). Originally designed to simulate paper-based tests, CKAT was not appropriate for many of the children with cerebral palsy in our user group because their impairment prevented use of the stylus on the tablet. CKAT was adapted to use the Microsoft gaming joystick as the interface between the child and the tablet or laptop in place of the stylus. The adapted CPKAT (known as CPKAT for use with children with cerebral palsy) was designed with tasks that required children to undertake three different types of tasks, all in the horizontal plane: fast, accurate aiming movements; steady, consistent tracking movements; and slow, accurate tracing tasks. These tasks are illustrated in Figure 2-12 on page 71.

However, CPKAT had not been tested or validated on children with cerebral palsy. As part of the study investigating the feasibility of school-based computer-assisted gaming technology described in sub-section 2.3.2 on page 66 a small sub-study was carried out (Preston et al., 2014a). The aim of the sub-study was to evaluate the feasibility of CPKAT as a kinematic assessment tool for use with children with cerebral palsy outside of the laboratory. It was hypothesised that it would be possible to distinguish between impaired and non-impaired arms of children with unilateral cerebral palsy using data captured by CPKAT for each arm. Conversely, it was hypothesised that no performance difference would be detected between the limbs on the tracking task (as reported within the unimpaired population) as unilaterally impaired children with cerebral palsy commonly experience visual disorders that would make this task difficult to complete e.g. predictive visual tracking, smooth pursuit, and with saccades (Salati et al., 2002).

3.6.1 Method

The first seven children recruited and assessed as part of the school-based feasibility study (2.3.2) performed the CPKAT assessment using their impaired arm (as part of the main study) and then repeated the exercise using their non-impaired arm. The conduct of the assessment and details of the tasks are described in Chapter 2, sub-section 2.3.2.5.3, pages 70-74, with the only difference being the use of the non-impaired arm after the initial assessment for the main study, for which the children used their impaired arm. The children's details are given in Table 3-23.

Table 3-23. Clinical and demographic details of children participating in the CPKAT trial.

Participant id	Gender	Age at assessment	Affected upper limb	Manual Ability Classification (MACS)	Gross Motor Function Classification (GMFCS)
1	Female	12 years	Right upper limb	IV	II
2	Male	8 years 8 months	Right upper limb	IV	II
3	Male	12 years 1 month	Right upper limb	II	II
4	Male	10 years 6 months	Left upper limb	II	II
5	Male	9 years 6 months	Right upper limb	III	II
6	Female	6 years 10 months	Right upper limb	III	II
7	Male	9 years 6 months	Right upper limb	III	II

3.6.2 Results

In the aiming task, CPKAT differentiated between the affected and non-affected side for Path Length Time (non-affected side was 21.9% quicker, $p = 0.028$) and smoothness (non-affected side showed positive difference of 58%, $p = 0.018$), but no differences were found for Path Length ($p = 0.237$).

No intra-limb differences in the tracking task were found between the affected and the non-affected side for any of the spatiotemporal parameters, either in the fast or slow tracking task.

In the tracing task, CPKAT differentiated between the affected and non-affected side for Path Length (non-affected side was 21.5% shorter p = 0.028) but not for Path Length Time (p = 0.398), Path Accuracy (p = 0.063) or TPA (p = 0.091). Detailed results are given in Table 3-24 below.

Table 3-24. Comparison of the upper limb kinematics of impaired and unimpaired arms of children with unilateral cerebral palsy.

TASK & KINEMATIC PARAMETER		MEDIAN	INTER QUARTILE RANGE	DIFFERENCE BETWEEN ARMS
AIMING TASK (PENTAGRAM)				
Path Length	Impaired arm	145.48	144.91 – 160.92	None, p=0.237
	Unimpaired arm	151.28	132.93 – 167.54	
Path Length Time	Impaired arm	2.55	2.43 – 4.06	Unimpaired arm 21.9% quicker, p=0.028
	Unimpaired arm	1.99	1.84 – 2.16	
Smoothness	Impaired arm	1256	1014 – 2738	Unimpaired arm 49% quicker, p=0.028
	Unimpaired arm	636	535 – 1070	
TRACKING TASK (SLOW FIGURE OF 8)				
Path Length	Impaired arm	1652	1584 - 1725	None, p=0.310
	Unimpaired arm	1595	1410 - 1667	
Smoothness	Impaired arm	129995	113463 - 150304	None, p=0.398
	Unimpaired arm	121040	81264 – 145533	

Accuracy	Impaired arm	18.5	13.30 – 24.75	None, p=0.610
	Unimpaired arm	19.63	11.37 – 26.63	

TRACKING TASK (FAST FIGURE OF 8)

Path Length	Impaired arm	3008	2770 – 3273	None, p=0.398
	Unimpaired arm	2754	2105 - 3180	
Smoothness	Impaired arm	108786	78941 – 167272	None, p=0.499
	Unimpaired arm	102709	65190 - 129132	
Accuracy	Impaired arm	29.33	25.46 – 43.37	None, p=0.176
	Unimpaired arm	32.64	17.22 – 34.45	

TRACING TASK

Path Length	Impaired arm	901	685 - 948	Unimpaired arm shorter, p=0.028
	Unimpaired arm	707	675 - 830	
Path Length Time	Impaired arm	41.55	21.6 – 48.89	None, p=0.398
	Unimpaired arm	30.01	19.75 – 35.84	
Path accuracy	Impaired arm	2.32	1.83 – 2.74	None, p=0.063
	Unimpaired arm	2.34	1.31 – 2.66	
TPA (path accuracy x Path Length Time)	Impaired arm	95.31	65.92 – 126.55	None, p=0.091
	Unimpaired arm	71.71	61.3 – 82.32	

3.6.3 Discussion

Repeated visits to the laboratory for kinematic evaluation were unreasonable for children with cerebral palsy and their families taking part in the school-based feasibility study. The RCT of the home-based computer-assisted arm rehabilitation games device following botulinum toxin treatment will recruit participants within a window of less than one hour, during which all assessments will have to be performed. It is therefore impossible for children to visit the laboratory for kinematic assessment in this hour before botulinum treatment. This feasibility study evaluated the use of a portable laptop-based system to evaluate upper limb kinematics of children with cerebral palsy in a non-laboratory setting. The parameters recorded by CPKAT were presentable for evaluation and statistical analysis. No problems were experienced by researchers or children, either in the transportation and set up of the CPKAT system, or in its use. All children reported enjoying the tasks.

Feasibility of using the portable system was therefore established but the main question now became of whether CPKAT could capture data that reveal differences in the children's upper limb kinematics. To answer this question, the kinematics for the impaired and unimpaired limbs of the children were compared. The results suggest that CPKAT measured statistically significant differences between the impaired and unimpaired limbs of children in both the aiming and the tracing task. In the aiming task, there were no intra-limb differences in path lengths suggesting that the path followed by each child was similar for each arm but the movement was faster and smoother for the non-affected arm. In the tracing task, there was a significant difference between the arms for path length, with the non-affected arm tracing a path over a fifth shorter than the affected arm. This might suggest that the tracing accuracy was significantly better but this was not the case, although it did approach significance. Path Length Time was no different between affected and unaffected arms. A composite measure that took both accuracy and movement duration

into account also approached significance. One possible explanation for this unexpected failure to detect differences of accuracy is tiredness and reduced concentration caused by the number of tasks and length of time taken by the tasks, particularly as the children performed them with each arm. For example, some children were still trying to do this final task quickly despite clear instructions to the contrary. These findings are consistent with previous reports within the literature of longer duration movements in the impaired arm of children with unilateral cerebral palsy (Jaspers et al., 2011, Ricken et al., 2005).

It was hypothesised that the limb differences would not be observed for the tracking task as the limiting factor for such tasks is the central ability to predicatively track the moving target with less demands made of the end effector (in contrast to the aiming and tracing tasks, in which the target is static). Inspection of the data showed that the CPKAT measures for tracking were similar between the two limbs. In short, the CPKAT system was capable of generating useful kinematic data from children with cerebral palsy in a non-laboratory setting.

These findings suggest that CPKAT can be employed in clinical trials where detailed kinematic data need to be collected in non-laboratory settings. There were no difficulties in understanding and executing the computer-orientated tasks themselves, however a number of issues arose when conducting the tests which may have affected the results. Two children were unable to achieve a hand grip on the joystick handle due to increased tone in the hand and arm, and did not complete any timed tasks fully. Successful capturing of their kinematic data was still achieved through an adapted grip on the joystick. The children all reported enjoying the tasks. In future, difficulties with maintaining grip could be minimised by modification of the interfacing joystick to account for increases in tone and reduced supination, or contractures causing fixed pronation.

One major advantage of the laboratory-based evaluation of upper limb kinematics is the capability to also monitor and evaluate trunk and shoulder movements. Evaluating upper limb movements using CPKAT does not account for shoulder and trunk movements. There are a number of ways of addressing this. Firstly, the CPKAT software and hardware includes a miniature inertial measurement unit (XSENS motion tracking technologies, Culver City, California), which measures 3-dimensional acceleration of the surface to which it is attached e.g. the shoulder. A wireless system is being developed that allows real-time monitoring and recording of the distance between the shoulder and the laptop during the tasks. This system can freeze the screen task if the user moves within a pre-defined distance, encouraging the user to refrain from using excessive trunk movement to compensate for restricted arm movement.

This study was conducted in a home setting using portable equipment, standard dining room tables and tables. There is no reason why CPKAT could not be used in the botulinum toxin clinic using standard office equipment, in paediatric establishments using child-appropriate seating and tables, or in any other clinical setting.

This study demonstrates that CPKAT has the potential to evaluate upper limb kinematics in children with cerebral palsy outside the laboratory setting. CPKAT is not designed to replace large lab-based kinematic measurement systems but rather complement them and provide a portable tool for monitoring and evaluating changes in upper limb kinematics in non-laboratory settings. Future plans include studies to test the psychometric reliability of CPKAT.

3.7 Conclusion of measures of activity limitation and kinematics

A detailed investigation into outcome measures for evaluating changes in upper limb activity limitation of children with cerebral palsy suggested that the AHA and the ABILHAND-kids are the most reliable options for therapists and researchers. For the purposes of the RCT to evaluate the benefits of the games device, the AHA was ruled out because of the extensive training and assessment procedure to validate therapists appropriately for its use. There are suggestions that the ABILHAND-kids is not appropriate for children with more severe activity limitation and doubts have been expressed about its responsiveness, but it was the most appropriate remaining option. The COPM was included for use in the trial because it is useful for developing individualised goals, and it is responsive; its limited use as a comparison across groups, however, is recognised.

The use of CPKAT to measure kinematics of the upper limb overcomes the obstacles associated with laboratory-based equipment. Adaptations of the CPKAT tests will be carried out to address potential limitations of tiredness and short concentrations spans.

A useful outcome of this phase of the study was the development of the ChARM, a new measure of activity limitation that potentially overcomes all psychometric limitations of each and every existing measure. However, the development of the ChARM was an unexpectedly lengthy process, so the ChARM was not available for use in the RCT.

Part 2 Randomised controlled trial to evaluate whether the use of a home-based computer-assisted arm rehabilitation games system enhances the functional benefits of botulinum toxin treatment of spasticity in the upper limb of children with cerebral palsy

4 Randomised controlled trial to evaluate whether the use of a home-based computer-assisted arm rehabilitation games system enhances the functional benefits of botulinum toxin treatment of spasticity in the upper limb of children with cerebral palsy

"knowledge cannot spring from experience alone but only from the comparison of the inventions of the intellect with observed fact"

~ Albert Einstein (1879 –1955)

"The true method of knowledge is experiment"

~ William Blake (1757 – 1827)

4.1 Introduction

Cerebral palsy is a common neurological condition of childhood, with a prevalence of up to 3.33 per 1000 births (Odding et al., 2006). Over 90% of children with cerebral palsy are affected by spasticity (Odding et al., 2006). This causes impaired movement, with upper limb activity limitation present in up to 83% of children with cerebral palsy (Odding et al., 2006). A number of other associated impairments may or may not be present (Rosenbaum et al., 2007, Novak et al., 2013), e.g. visual impairment (up to 71% of children with cerebral palsy; Odding et al., 2006) and learning disability (up to 60% of children with unilateral impairment; Odding et al., 2006).

Since the beginning of this century, upper limb activity limitation has been the focus of experimental rehabilitation programmes based on theories of motor learning. Studies into Constraint Induced Movement Therapy (CIMT) and bilateral training therapies have produced moderate evidence to support the proposals that repetitive and intensive practice of unilateral and bilateral activity promotes functional independence and reduces activity limitation in children with cerebral palsy (Gordon et al., 2007, Green et al., 2013, Hoare et al., 2007a, Huang et al., 2009, Sakzewski et al., 2011b, Sakzewski et al., 2009). Intensity of practice and number of repetitions appear to be the essential component. There are suggestions that a programme of unilateral training (CIMT) followed by goal-directed and other bilateral training that stimulates high intensity and repetition might be the most effective pathway (Aarts et al., 2010, Andersen et al., 2013), but no studies have investigated this. There is strong evidence that the use of botulinum toxin as a targeted treatment for spasticity has a beneficial effect on daily activity when combined with

rehabilitation therapy (Hoare et al., 2010), and studies investigating the combination of CIMT and bilateral training with botulinum toxin treatment have shown promising results (Hoare et al., 2013).

The intense nature of CIMT and bimanual training appears challenging to some children and families, but growing interest in the use of virtual reality, video-gaming and robotics to supplement rehabilitation programmes might lead to greater adherence and participation in intensive rehabilitation programmes. However, studies into the use of these technologies are limited mainly to those that produce low-level evidence such as case studies (Krebs et al., 2009, Meyer-Heim and van Hedel, 2013, Sandlund et al., 2009). Feasibility studies of computer-assisted arm rehabilitation (CAAR) gaming technology were carried out, initially in children's homes and then, following a period of redesign and development, in children's schools. These studies provided limited support for the potential of assistive gaming technology to reduce upper limb activity limitation and demonstrated improved upper limb kinematics of children with cerebral palsy. Low-level evidence also suggests that robotics and virtual reality used after botulinum toxin treatment for spasticity could result in the same functional benefits as an occupational therapy programme when that is combined with botulinum toxin (Fasoli et al., 2008).

In recognition of the potential additional benefits of using assistive computer technology in combination with botulinum toxin treatment of the upper limb, it was proposed to carry out a research study to investigate the benefits of using the CAAR games device in combination with botulinum toxin treatment.

4.2 Method

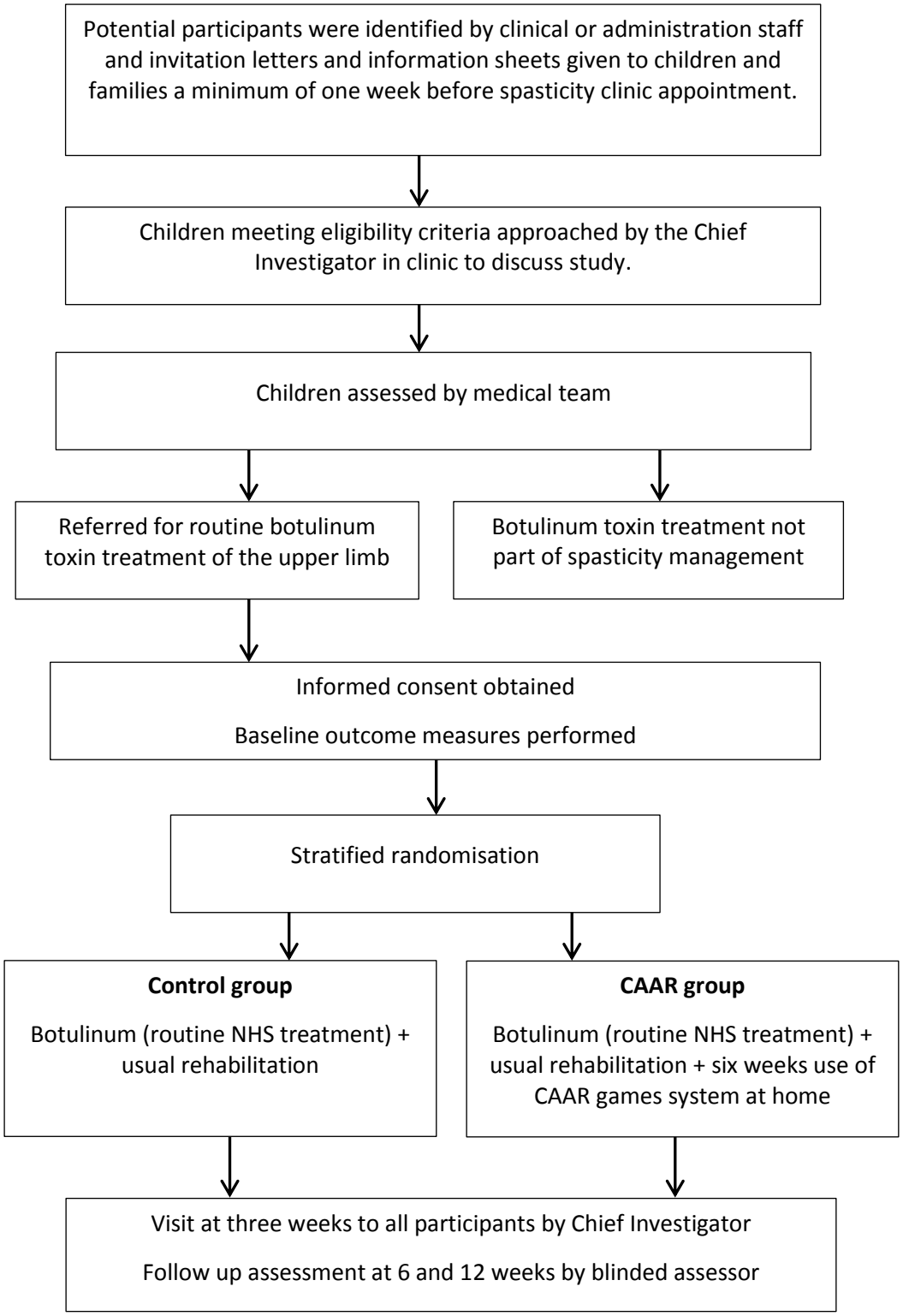
4.2.1 Study design

The best design for an efficacy study, such as a study to investigate whether use of the home-based CAAR game enhances activity limitation benefits of botulinum toxin treatment for spasticity of the upper limb of children with cerebral palsy, is a randomised controlled trial (RCT; Altman, 1996, Schulz et al., 2010). Following the framework established by the Medical Research Council (MRC), this study was a Phase II exploratory study (Campbell et al., 2000). The study was essentially a feasibility and pilot stage (Craig et al., 2008) to prepare for a Phase III multicentre study.

A useful methodological inclusion to an RCT (and other types of study) is blinding (Day and Altman, 2000). Blinding is used to limit the influence of human preconceptions, called bias

(Day and Altman, 2000). In a double-blind trial neither the participants nor the investigator are aware of the allocation of the participant (Schulz and Grimes, 2002). In this study it was impossible to mask the participant to the allocation i.e. whether they were playing the CAAR games device or not, but of particular importance when this is not possible is blinding of the assessors (Schulz and Grimes, 2002). The design in this case was a single-blind design, in which the baseline and follow up outcome measures were carried out by assessors blinded to the allocation. The conduct of the study is outlined in Figure 4-1 and is described using the CONSORT guidelines (Craig et al., 2008) for parallel group randomised trials (Schulz et al., 2010).

Figure 4-1. Flow diagram outlining study design.



Potential participants were children with cerebral palsy who were referred to one of four regional spasticity clinics and who met the eligibility criteria in Table 4-1.

Table 4-1. RCT eligibility criteria.

Inclusion criteria	Exclusion criteria
Children aged 5 – 12 years old with diagnosis of cerebral palsy	Children who have had upper limb surgery within the previous six months
Children referred for treatment with botulinum toxin for upper limb spasticity	
Children with MACS Levels II-IV	
Sufficient cognitive ability to play simple computer games	
Able to manipulate handle of robotic arm	
Vision sufficient to view computer screen and follow on-screen movements	

MACS: Manual Ability Classification System

4.2.2 Funding, sponsorship, trial registration, regulatory approval and National Institute for Health Research (NIHR) support

4.2.2.1 Funding

The study was funded by the NIHR as an educational grant in the form of an NIHR Clinical Doctoral Fellowship.

4.2.2.2 Sponsor

The study was sponsored by the University of Leeds.

4.2.2.3 Trial registration

The RCT was registered on the International Standard Randomised Controlled Trial Number Register, and was allocated number ISRCTN26206379 (http://www.controlled-trials.com/ISRCTN26206379/).

4.2.2.4 Ethical approval

Ethical favourable opinion was given on 23rd September 2011 by NRES committee Yorkshire and the Humber – Leeds West (REC reference 11/YH/0276).

4.2.2.5 Medicine and Healthcare products Regulatory Agency (MHRA)

There were a number of issues that had to be addressed before the MHRA were prepared to publish a letter of no objection.

1. It was recognised that the children allocated to the control group (the group not receiving the CAAR games device) might be disappointed not to have opportunity to play the CAAR games system. To alleviate this disappointment, and avoid the possibility that allocation to the non-games group might precipitate withdrawal of the child from the study, the study protocol included a week of CAAR device use for the control group children, at home, after the children had concluded their part in the trial. This was unacceptable to the MHRA, and the offer of the opportunity for the control group to play the games system was removed from the protocol.
2. A panel was to be attached to the top of the CAAR device warning non-participants e.g. siblings to refrain from using the CAAR device.
3. Changes to the Patient Information Sheet and the User Operating Instructions were requested:
    a. A prominent note was added to the Patient Information Sheet to make clear that use of the CAAR was to be supervised by parents.
    b. A warning that only study participants were to use the CAAR device was added to the User Operating Instructions.

A letter of no objection was received from the MHRA on 12th March 2012 (MHRA reference CI/2012/0005).

4.2.2.6 Adoption by the NIHR

This study was adopted by the NIHR on 26 September 2011. Adoption onto the NIHR portfolio allows studies access to the Coordinating System for obtaining NHS Permission (CSP).

4.2.3 Approaching and recruiting participants

Four regional spasticity clinics supported the approach and recruitment of children to the RCT:

- Leeds Teaching Hospitals NHS Trust (LTHT);
- York Teaching Hospitals NHS Trust;
- Huddersfield and Calderdale NHS Trust;
- Hull and East Yorkshire NHS Trust.

The approach and recruitment of children was different at the LTHT spasticity clinic to the other three sites but was designed to maintain patient confidentiality throughout, until the child and the family notified the clinic staff that they were interested in discussing the study with the Chief Investigator.

The LTHT holds spasticity clinics every week, each with at least two consultants to assess children and administer botulinum toxin injections if clinically appropriate. In addition, there are irregular orthopaedic clinics conducted by an orthopaedic surgeon who offered to support recruitment to the study. The other three sites held clinics approximately once a month; these were conducted by a single consultant to administer botulinum toxin treatment. It was anticipated that LTHT would recruit a greater proportion of the intended sample; however, it was anticipated that a minimum of ten participants would be recruited by the other three sites over two years.

4.2.3.1 Approach to participants at the Leeds site

The LTHT regional spasticity clinic's clinical lead was an experienced paediatric consultant neurologist, whose secretary identified appropriate potential participants from the clinic list. The secretary then posted information sheets and invitation letters to the children and their parents, and notified the Chief Investigator of the clinics at which potential participants would be in attendance. At the clinic, medical staff were able to direct the family to the Chief Investigator for a discussion about the study. If parents and children expressed an interest in participating in the study, inclusion and exclusion criteria were assessed.

The LTHT regional spasticity clinic is both a botulinum toxin assessment and treatment clinic. Children attending the clinic were first assessed by a multi-disciplinary team including neurology consultants, occupational therapists and physiotherapists to evaluate whether the clinical and functional benefits of treating the child with botulinum toxin. If appropriate, children were then treated with botulinum toxin within the same clinic, usually within an hour. If the family had indicated their willingness to participate, the decision to administer botulinum toxin fulfilled the final criterion for inclusion and the process of taking written informed consent and performing baseline assessments took place prior to botulinum toxin administration.

On several occasions, an appointment was cancelled by potential participants and their place at the clinic taken by another child that met eligibility criteria. Under the terms of initial ethical approval, these children could not be approached because they had not received information sheets a week before the appointment. This was addressed through a

protocol amendment that permitted an approach to these families, given that this non-invasive study was unlikely to result in harm to the child. Families had the opportunity to take the information sheets away to review them at their leisure in the days after the clinic appointment, and had the opportunity to withdraw from the study before taking delivery of the CAAR device. If the participant were to withdraw in circumstances such as these, the participant would be removed from all trial procedures and programs as if they had refused initially to participate.

The paediatric orthopaedic service for botulinum toxin treatment included two clinics. Children were assessed by the orthopaedic surgeon and the attending occupational therapist at a general orthopaedic clinic. One possible outcome of this assessment clinic was a referral to the orthopaedic botulinum clinic which was held a few days later. Potential participants were informed of the study and given the appropriate information sheets. The Chief Investigator was usually available to attend the assessment clinic at short notice if an eligible child was identified but if he was not available to attend the assessment clinic he attended instead the botulinum toxin clinic to discuss the study and to take informed consent and baseline measures.

4.2.3.2 Approach to participants at sites other than Leeds

At the other three sites, assessments for botulinum toxin treatment and administration of botulinum toxin were performed at different clinics. Initially, children attended an assessment clinic at which the clinical decision was taken about whether the child would be referred for botulinum toxin treatment. The assessment clinics were staffed by experienced paediatric therapists who directly supported the study though the identification of eligible children. If referred for botulinum toxin, the treatment clinic appointment was within one to three weeks of the assessment clinic. Children referred for botulinum toxin treatment who met eligibility criteria were given information sheets by clinic therapists who informed the Chief Investigator of the forthcoming appointment. The Chief Investigator attended the botulinum toxin treatment clinic to discuss the study with the family, take informed consent and carry out baseline measures.

4.2.3.3 Recruitment: informed consent

If children and their families agreed to take part in the study, written informed consent was obtained from parents and children aged 12 years and over. Written assent was taken from children under 12 years old. This procedure took approximately 10 to 15 minutes.

4.2.4 Sample size

As this was a Phase II study and due to the lack of published data regarding effect size sample sizes were calculated using published data about psychometric testing of the primary outcome measure, the ABILHAND-kids. Using this information, the sample size calculations suggested that it would be necessary to recruit 58 children (29 patients per group) in order to detect a large effect at 5% level of significance with 80% power. This number of participants would also allow us collect information about feasibility and trial fidelity issues in a future larger scale study. This information would include barriers and enablers to device deployment, how much the device was used and the occurrence of any safety issues while the device was being used in children's homes.

4.2.5 Baseline assessments

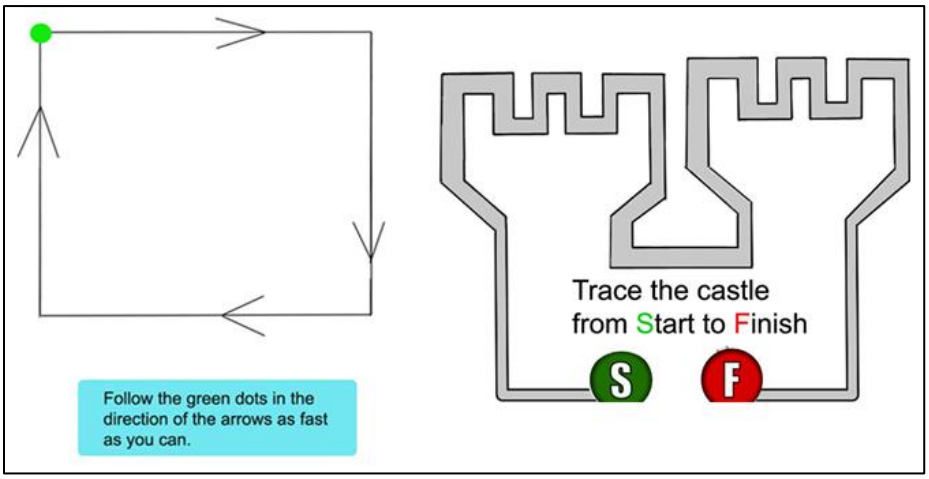
Baseline assessments were performed before treatment with botulinum toxin. As discussed at length in Chapter 3 and summarised in sub-section 3.3.5 on page 141, the outcome measures selected for the evaluation of upper limb activity were the ABILHAND-kids (primary outcome measure) and the Canadian Occupational Performance Measure (COPM). These are described in detail in Table 3-4 and Table 3-5 in Chapter 3, subsection 3.3. The ABILHAND-kids is a questionnaire of 21 items. The 21 items are presented on ten sheets, each with the same 21 items but in a different order. This is designed to prevent parent recall of answers and overcome response bias. For the baseline assessment, sheet 1 was used for each participant. The CPKAT was used for kinematic assessments. The child was directly involved in the kinematic assessment only, which took approximately ten minutes. Only the child's parent was involved with the activity measures. The assessments took approximately 20 minutes altogether.

The CPKAT tasks that were described in subsection 2.3.2.5.3 and validated as described in subsection 3.6 were modified so that there was less chance of the children becoming bored, fatigued or losing concentration. The Figure of Eight tracking tasks described previously in sub-section 2.3.2.5.3 and shown in Figure 2-12 on page 71 were removed. The Pentagram was retained. The tracing shapes were changed and reduced in number from four to two tracing tasks. These omissions and changes did not limit the kinematic parameters captured by CPKAT.

The kinematic practice tasks were also modified and are illustrated in Figure 4-2. Children first practiced on the Square aiming task and then practiced on the Castle tracing task. These were practices for the Pentagram aiming assessment and the House tracing assessment

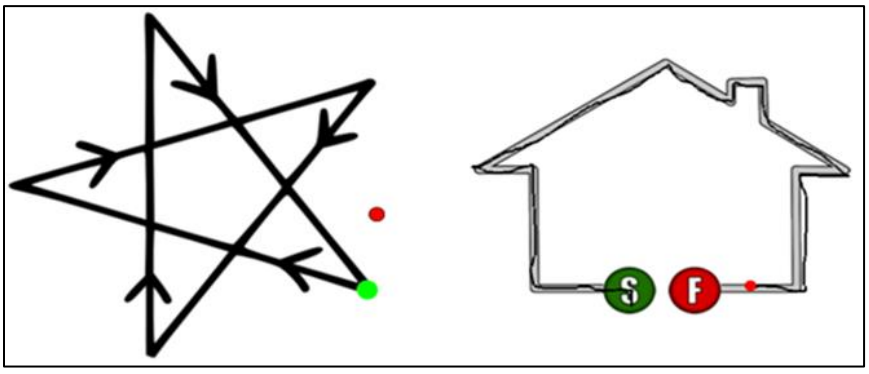
respectively. The Square practice lasted 30 seconds and was performed twice; the Castle practice was untimed and children were encouraged to take as long as they needed to draw an accurate trace. For this reason, the practice was only performed once to prevent tiredness and loss of interest.

Figure 4-2. The practice tasks for the CPKAT assessments: the Square aiming practice (left) was practiced twice andeach lasted 30 seconds; the Castle tracing practice (right) was untimed and was practiced once.



The assessment tasks are shown in Figure 4-3.

Figure 4-3. The assessment tasks. The Pentagram aiming assessment task (left) was performed twice and each lasted 30 seconds. The House tracing assessment task (right) was untimed, and was performed twice.



The Pentagram assessment task lasted 30 seconds and was performed twice, after all practices on the Square and Castle had been completed. The House tracing task was untimed and children were asked to take their time and to be as accurate as possible. This task was also performed twice. Children were told that the best score of their two attempts at the assessment tasks would be used for their final scores. The reason for selecting the best score over a mean score was to encourage the child to concentrate more fully

throughout the tasks, and to try harder for their second attempt. This would address the tendency in the feasibility study for children to explore and experiment with the tasks. It was also considered that the best score would more accurately capture the child's ability.

Because baseline assessments were carried out before randomisation, the assessor was blinded to allocation (Day and Altman, 2000).

4.2.6 Randomisation (stratification using minimisation)

Once baseline measurements were completed, randomisation was carried out. Randomisation is used to reduce unsystematic variation between experimental groups (Field, 2009) and to limit bias by preventing the influence of investigators on allocation of participants to the experimental groups (Altman and Bland, 1999b). Such bias might be the allocation, for example, of a participant to the non-intervention group (even if sub-consciously) because of characteristics which cause the allocator to believe that the participant would not show an improvement in the intervention group (Altman and Bland, 1999b). Unsystematic variation is due to the natural differences in characteristics between people (Field, 2009) e.g. a personal preference for playing computer games.

To be able to draw conclusions from the comparison of one group of people to another, characteristics e.g. gender, manual ability, age should be balanced across the groups. In small studies, group sizes are too small for natural balance to occur through random allocation. In this case, stratified randomisation is used to balance groups without compromising the random process of allocation (Altman and Bland, 1999a). A further advantage of stratification is that it can achieve balance in specific characteristics that might influence the outcome of the study (Altman and Bland, 2005, Altman and Bland, 1999a). These characteristics are pre-determined by investigators e.g. younger children (aged 5 – 7 years) are likely to have much less time available to play the CAAR system due to bed time schedules so that unbalancing of the groups in these ages might influence the outcomes.

The groups in this RCT were stratified using a technique called minimisation (Altman and Bland, 2005). Minimisation balances the factors shown in Table 4-2, which were considered to have the potential to influence outcomes. Minimisation is a useful alternative to stratified randomisation, especially in small trials when there are strong prognostic factors that potentially will influence outcomes and when modest treatment effects are anticipated (Altman and Bland, 2005).

Table 4-2. Stratification factors.

Factors	Age (three strata)	Gender (two strata)	Home computer games (two strata)	MACS (three strata)
	5 – 7 year olds	Male	Yes	MACS Level II
Strata	8 – 10 year olds	Female	No	MACS Level III
	11 – 12 year olds			MACS Level IV

Minimisation ensures that groups remain balanced on stratifying factors by summing a participant's stratified factors for each experimental group every time a participant is recruited, but before the participant is added to a group. The participant is then added to the group with the lowest total of summed stratifying factors, each of which are incremented by one in that group. An example from the study is given in Table 4-3 below. This shows the stratified factors of the study participants after seven participants have been recruited. Details of the eighth participant's factors are in the second column to aid understanding of the summing process: a female aged between five and seven years old, MACS level IV and who has computer games at home. These factors are summed for each group, and the eighth participant is allocated into the group with the smallest total.

Table 4-3. An example of minimisation, illustrating how the eighth participant was allocated.

		CAAR device group		Control group	
	New (eighth) participant strata	Current factor totals	Summed factors	Current factor totals	Summed factors
Aged 5 - 7	✓	0	0	1	1
Aged 8 - 10		3		2	
Aged 11 - 12		1		0	
Male		2		2	
Female	✓	2	2	1	1
MACS Level II		1		0	
MACS Level III		2		1	
MACS Level IV	✓	1	1	2	2
Computer games at home	✓	3	3	3	3
No computer games at home		1		0	
TOTAL			6		7

Summed intervention factors matching the new participant total less than the summed control factors, therefore the participant is allocated the CAAR device group.

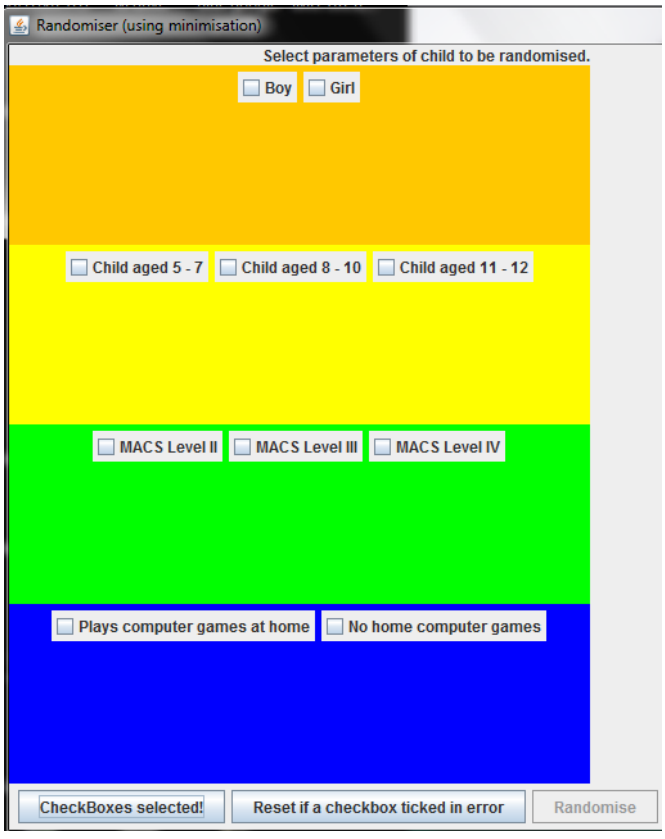
If summed factors are equal, the participant is allocated to a group using simple randomisation.

There are three main problems with minimisation (Scott et al., 2002): the potential for errors with summing multiple factors; the possibility that each new participant's allocation can be predicted, thus potentially introducing bias; and the use of non-random allocation which undermines the assumptions necessary for some statistical analyses. There is also the possibility of a participant withdrawing, thus unbalancing the groups and causing all participants minimised since the withdrawn participant's entry to have been allocated based on incorrect summed factors.

Scott et al. (2002) review a number of authors' views on the subject of compromising statistical assumptions and introducing wide confidence intervals on analyses through the lack of true random allocation. There is no consensus but most authors do not believe that minimisation has a negative influence on the statistical outcomes, as long as the minimisation factors are included as variables (covariates) in an analysis of covariance. The review suggests that the minimisation procedure increases the statistical power of an unadjusted analysis because there are fewer errors with unbalanced groups, and points out that the potential still exists for each participant to be allocated to any group. This potential can be increased by introducing a random weighting into the allocation process that allocates the participant to the group that *further unbalances the group* (Altman and Bland, 2005). Another problem is the difficulty in using minimisation in multi-site studies, with each local investigator needing up-to-date details of the participants' characteristics.

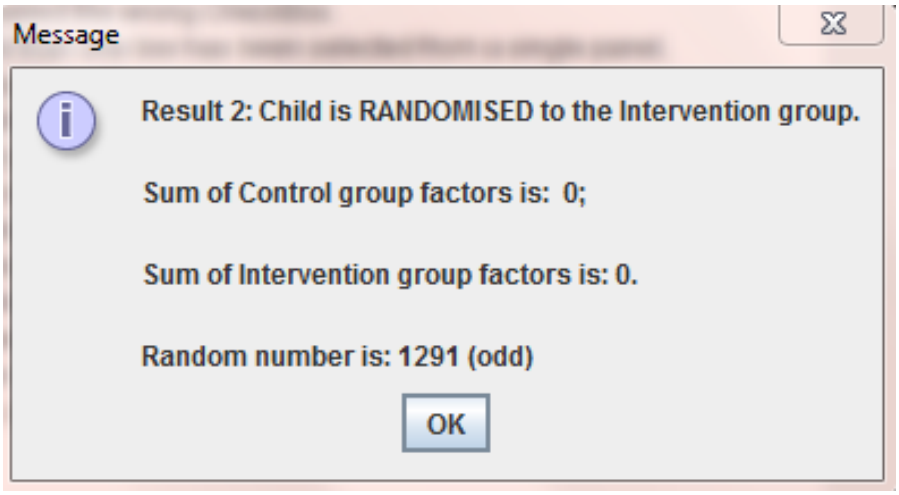
To try to overcome these problems, the minimisation procedure was performed by a bespoke computer program. Written in Java programming code to ensure compatibility on most platforms (and within web browsers), the program is easy and quick to use. The researcher enters the appropriate stratifying factor in the Graphical User Interface (GUI) shown in Figure 4-4, and presses the 'Randomise' button.

Figure 4-4. The Graphic User Interface (GUI) for the minimisation program.



The allocation resulting from the minimisation process is shown by the window illustrated in Figure 4-5, along with other information that is useful for checking the correct operation of the minimisation process.

Figure 4-5. Minimisation program window showing result of minimisation process.



The window indicates the group to which the participant is allocated. The information includes the totals of the summed factors and a program-generated random number. The example shows the first participant allocated to a study, so summed factors are nil in each group and the participant is randomly allocated - in this case, to the intervention (CAAR) group. Random allocation depends on the parity of the random number, with odd numbers allocated to the intervention (CAAR) group and even numbers allocated to the control group. On closing down the windows, the program saves a file for each group to the program folders with updated factors, in preparation for the next participant to be minimised. Using this program, therefore, removes the potential for summing errors. It also removes the introduction of bias, because there is no requirement for any researcher to see the previous participants' prognostic characteristics on which the factors are summed, although it does not remove the potential for intentional manipulation of the groups' allocation.

Concealment of randomisation procedures is recommended for limiting subconsciously or purposefully introduced bias (Altman and Schulz, 2001). To achieve this, the minimisation procedure was carried out by an independent third party: clerical staff within the regional community paediatric physiotherapy team. The program code and files were inaccessible to any research staff once installed on the community team's shared drive. Clerical staff had no knowledge of the participants except for the stratifying factors which were given during the minimisation procedure. The minimisation procedure was carried out remotely by telephone. This removed the potential for manipulating the groups' allocations. However,

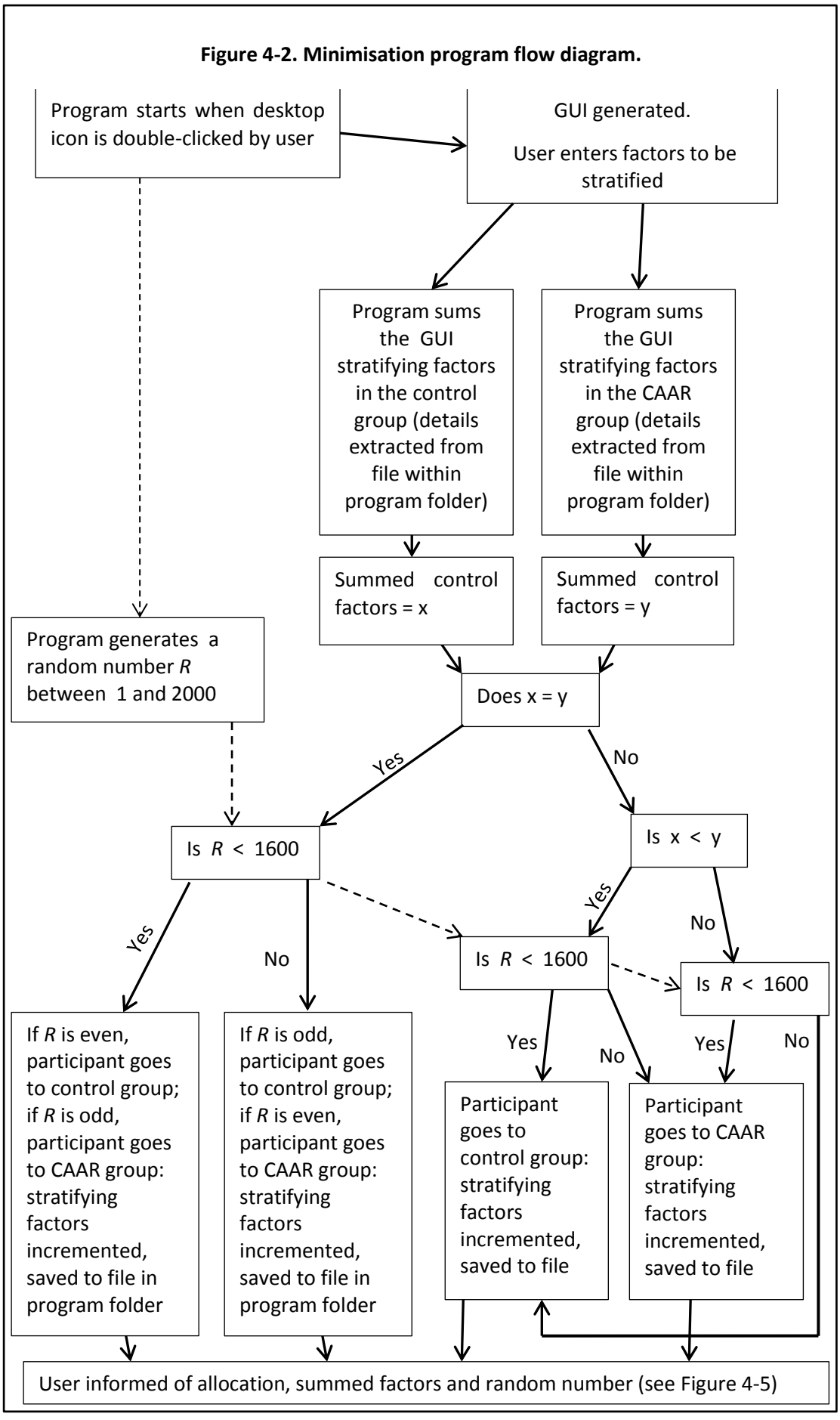
clerical staff sickness and increased security restrictions on the regional team's computer network proved insurmountable obstacles for them to be able to access and run the program, and ended their involvement in the minimisation process after the allocation of the first ten participants. The minimisation program was subsequently installed instead on the CPKAT laptop and minimisation was carried out by the Chief Investigator.

Children in this study were allocated to groups in a 1:1 ratio. If necessary, the minimisation programme could be modified to randomly allocate study participants in any ratio that investigators decide *a priori.*

A facility was included within the program to remove any participant that withdrew from the RCT.

Finally, the program included a random element that allocated the participant to the group that *further unbalanced the* group, as suggested by a number of authors (Altman and Bland, 2005, Scott et al., 2002). This was arranged so that when the minimisation program had completed its summing of factors and had calculated the allocation of the participant, its final action was to check the random number. If the random number was greater than 1600, it allocated the participant to the opposite group to that which reduced the unbalancing of the groups. Because the random number generated was between 1 and 2000, this was a probability of 0.2 that the groups would be further unbalanced. The minimisation program flow diagram illustrating how the program operates is shown in Figure 4-6.

Figure 4-2. Minimisation program flow diagram.



Parents and children were informed of the allocation to the CAAR group or the control group following the minimisation procedure, and arrangements made for follow-up as appropriate for each group. The blinded assessor who was to perform follow-up assessments at six and twelve weeks was given dates six weeks and twelve weeks from the first Monday of the week after treatment. This was to reduce the chance of the blinded assessor guessing the allocation of children through the dates of the follow up e.g. six weeks after treatment for the control group and seven weeks for the CAAR group, but still allowed a full six weeks of games use.

4.2.7 Control group and CAAR group

Both groups were treated in the same way apart from allocation of the CAAR games device to the CAAR group. Each group had botulinum treatment following clinical assessment by the medical team. Referral to the spasticity clinic, assessment and treatment were all independent of the RCT.

Follow up by paediatric therapy teams was also independent of the RCT. Potentially, some children could receive intensive therapy from their local rehabilitation team in support of the botulinum treatment, while others might receive no therapy. Although the randomisation process should balance this, the methodology attempted to control for it by providing parents with weekly diary sheets (see Figure 4-7) to indicate daily rehabilitation and other potentially beneficial activity that could be incorporated into statistical analyses.

Figure 4-7. Weekly diary for parents to record daily activity that might influence additional improvements to activity limitation.

Day	Therapy (occupational therapy, physiotherapy or other)	Use of other computer games	Any other exercise or unusual event relating to arm exercise
Saturday	e.g. daily stretches, daily functional activity to practice, repetitive strength or movement exercises.	e.g. Wii, Xbox, PC, use of joystick etc	e.g. swimming, wheelchair basketball.
Sunday	Therapy type: Time spent (minutes):	How long played (minutes): Type of game played:	How long child exercised (minutes): Type of exercise:
Monday	Therapy type: Time spent (minutes):	How long played (minutes): Type of game played:	How long child exercised (minutes): Type of exercise:
	Therapy type:	How long played (minutes):	How long child exercised (minutes):

4.2.7.1 Delivery and use of the CAAR device

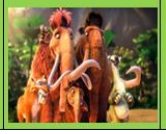
The CAAR device was delivered within one week of botulinum toxin treatment and collected six weeks after delivery. The six week period was chosen to take advantage of the optimal period of effectiveness of botulinum toxin (Yang et al., 2003) and to ensure sufficient quantity of practice and repetition. The discussion of CIMT and bilateral training in sub-

section 1.6.4.2 on page 40 suggests that the optimal dose of these therapies has not been established (Andersen et al., 2013), and there is a wide variation in quantity of rehabilitation using these methods. Thirty hours (Aarts et al., 2010, Hoare et al., 2013, Sakzewski et al., 2011b) is unrealistic for children attending school, however, and benefits were reported in studies using much less active rehabilitation (Eliasson et al., 2005, Wallen et al., 2011). Parents were therefore asked to allow children to use the CAAR device for 30 minutes a day, but it was suggested to parents that limiting use of the CAAR to periods of ten minute periods during the first week might help prevent muscle fatigue and aching due to unaccustomed exercise.

Delivery and collection were two-person tasks, and were arranged for the convenience of the family. Once the CAAR device was installed within the home, the parent(s) were shown how to operate the device. An instruction manual was included with the delivery, and device operation was explained to the parents in line with the instruction manual. The games system was tested in situ, with the parents practicing device set up.

Based on the experience gained in the first home-based feasibility study, it was recognised that parental support and encouragement for the children was essential if the required intensity and quantity of games use was to be achieved. To encourage parents' active support stickers were issued along with the chart shown in Figure 4-8 to help them to support and encourage their child in the periods when their enthusiasm waned.

Figure 4-8. Chart issued to parents along with stickers for parents to encourage daily CAAR device playing of set periods.

Monday	Tuesday 	Wednesday	Thursday	Friday	Saturday	Sunday 
Monday	Tuesday	Wednesday	Thursday 	Friday	Saturday	Sunday
Monday	Tuesday	Wednesday 	Thursday	Friday	Saturday 	Sunday
Monday 	Tuesday	Wednesday	Thursday	Friday 	Saturday	Sunday

4.2.8 Maintenance and check visit

A visit to check the CAAR at the participants' home was carried out after three weeks. The purpose of the visit was to offer encouragement to the children and to check the CAAR system. To maintain balance between the groups, a visit was also carried out to the control group. The control group children were offered similar encouragement to participate in any rehabilitation program that they had been given by their rehabilitation team or therapists at the spasticity clinic.

4.2.9 Collection of CAAR games device

The device was collected as close to the date on which the blinded assessor was to arrange their visit for the six weeks follow up assessment. This was given as six weeks from the first Monday after treatment.

Prior to collection, a letter was sent to the family requesting that they completed their diary describing rehabilitation performed daily, use of commercial computer games and any other activities. The letter included a questionnaire to gather feedback about use of the games system, their engagement with the games and about their participation in the study. A pre-addressed postage-paid envelope was included in which families were requested to place

the documents, either to post them, or hand them to the blinded assessor on their visit. This standard procedure was in place to prevent the blinded assessor identifying which houses had been visited for the purpose of collecting the device.

Parents and children were reminded in the week before the blinded assessor visit that all details of their participation in the study were to be kept secret from the blinded assessor.

4.2.10 Selection and training of the blinded assessor

An independent assessor who was unconnected and previously unknown to the research team was recruited for the purpose of carrying out blinded assessments at six and twelve week time points. Initially, regional physiotherapy schools were approached with the intention of recruiting a student physiotherapist. When no interest was received through these approaches, an advertisement was posted on the interactive Chartered Society of Physiotherapy website forum for 'Newly qualified physiotherapists'. Two physiotherapists responded: one based in London and the other travelling overseas. Finally, a poster was emailed to second year students within the Institute of Psychological Sciences, School of Medicine and Health, University of Leeds. Four Psychology students responded. They were all invited for interview by the Chief Investigator. Two failed to arrive for interview, but the two attending interviewees were well-prepared and enthusiastic.

Selection of the blinded assessor was made based on experience of working with disabled children, and better preparation and content of the ten minute presentation that was required of the applicants. As a second year student, the blinded assessor was expected to be available for the lifetime of the study, thus ensuring continuity and consistency of follow up assessments. The blinded assessor was reimbursed for travel expenses and for any hours spent on duties connected with the RCT.

The blinded assessor was given training in the CPKAT device and issued with a laptop and joystick. Although it was essential that baseline assessments using the COPM were performed by an experienced therapist (the Chief Investigator), the follow up assessments for the COPM and the ABILHAND-kids were more straightforward. To further prepare the blinded assessor for independent follow up assessments, the blinded assessor attended the first clinics in Leeds at which recruitment was anticipated to take place, and was able to observe the baseline assessments. Subsequently, in clinics at which more than two potential participants were attending, the blinded assessor was able to attend to carry out CPKAT assessments. This increased the amount of time available for the Chief Investigator to take informed consent and carry out the initial COPM and ABILHAND-kids measures.

4.2.11 Six week and twelve week follow up assessments

The blinded assessor arranged visits to participants' homes at six and twelve week to perform follow up assessments as close as possible to the dates provided by the Chief Investigator after randomisation.

At the six week assessment the ABILHAND-kids sheet provided was sheet order 2. No record or reminders were given of the responses at the baseline assessment, which was completed on sheet order 1, thus avoiding the possibility of response bias. The 12 week assessment in turn used sheet order 3. This was the same for all participants to reduce the potential that completing items in different orders between participants could influence the outcome score.

The COPM goals were provided on a form prepared by myself, without the scores given by parents at the baseline assessment. For the majority of the participants, the blinded assessor had no knowledge of the responses to COPM items given at baseline; it is unlikely that the responses could have been recalled, but the blinded assessor was instructed not to guide the parents in their responses to the items. The forms were retained by the blinded assessor between the six and twelve assessments so the blinded assessor was instructed again not to reveal the six week responses for either outcome measure, nor to guide parents.

The form included questions to be completed by the blinded assessor, about whether the blinded assessor could identify the group into which the participant had been allocated. These questions were designed to evaluate the success of the blinding strategy.

4.3 Statistical analyses

All statistical analyses were performed using IBM® SPSS® Statistics Version 21 Release 21.0.0.0 64 bit edition.

Primary analysis was on an 'Intention to treat' basis, with statistical significance assessed at the 5% level. The outcome analyses assumed that children lost to follow-up did not show any improvement. The ABILHAND-kids raw scores were transformed into interval level data and, if appropriate, parametric statistical analyses were performed on these scores and kinematic outcome scores using general linear models (ANCOVA, regression; Field, 2009), adjusting for child covariates (age group, gender, use of other computer games and MACS levels). Non-parametric analysis was used if the data did not meet requirements for parametric analysis. The ANCOVA permits multiple t-tests to determine whether there are

significant differences between groups at each time point and within groups across time points, adjusting for multiple testing in order to prevent a Type I error (Field, 2009, Pallant, 2007). The ANCOVA also adjusts for additional variables (covariates), the inclusion of which was necessary to counter arguments against the use of minimisation, as discussed in subsection 4.2.6 (page 216). If the ANCOVA detected significant differences between groups or time points, an exploration of differences was carried out using multiple regression. This included an investigation into the effects of the prognostic factors and additional covariates such as amount of time played, age, gender and MACS levels.

COPM outcome scores are ordinal data. COPM results were analysed using non-parametric ANOVA (Friedman's ANOVA and Kruskal-Wallis Test for within-groups and between-groups analyses respectively). If significant differences were found in the repeated analyses or between-groups tests, further evaluations were performed using appropriate *post hoc* testing.

The two tasks of CPKAT evaluate a number of different parameters which produce outcome scores of interval level data. Once assumptions of Normality were checked, the appropriate parametric or non-parametric ANOVA was used to investigate any differences between time points (within-group changes) and between-groups at each time point. Appropriate *post hoc* tests were performed if differences were indicated.

The success of the blinding strategy was evaluated using a binomial test of proportions.

4.4 Results

4.4.1 Recruitment, consent and study profile

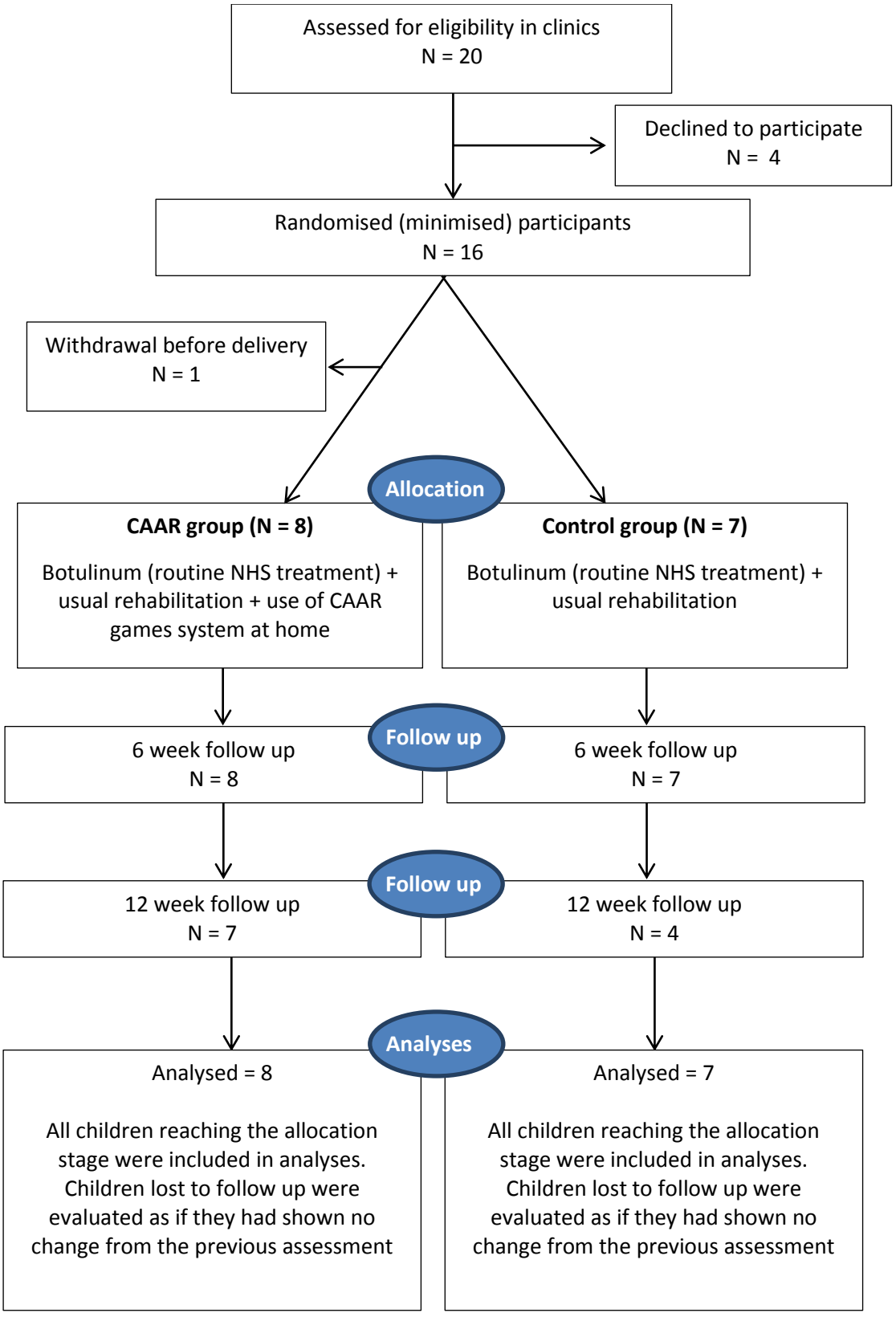
The trial profile (Schulz et al., 2010) is shown in Figure 4-9 below. Twenty children were identified as potentially eligible and approached for participation in the study. Sixteen children from three participating sites were allocated to the two groups through the minimisation program over a period of 21 months. One parent withdrew the child from the study, after they had given informed consent and after the child had been allocated to a group, but before taking delivery of the CAAR device, citing lack of room for the device. This participant was not included in any statistical analyses, and their stratifying factors were removed from the minimisation CAAR file using the built-in functionality of the minimisation program.

Four parents of eligible children refused to participate in the study. One parent explained that the device was too large for installation in their very small home, and refused with regret. The second parent refused on the principle that it was inappropriate and harsh to present a games system to a child and then withdraw it, suggesting that this would cause the child distress. The third parent refused to participate because they believed that their child would not be able to play the games even with the assistance of the robotic arm and even if botulinum toxin produced any functional benefits. The fourth parent was approached at the LTHT orthopaedic clinic. The parent expressed an interest at the initial meeting but refused permission for contact details to be given to the research team and did not follow up on their offer to contact the research team and confirm their participation.

Four children were lost to follow (LTF). One parent reported that they were too busy to meet for the final assessment at twelve weeks, two parents were inaccessible for all attempts to contact them, and one child was recuperating from elective surgery on a pre-existing medical condition that was made known to me prior to their recruitment and informed consent procedure.

All other participants were allocated to a group and included in all statistical analyses.

Figure 4-3. Trial profile for consent, participation and follow up.



4.4.2 Record of time lines

Table 4-4 shows the number of days between trial procedures. Informed consent, baseline assessments and botulinum treatment all took place on the first day.

All CAAR devices were installed within seven days of botulinum treatment. The CAAR games device was installed for use in children's homes for a mean of 40 days (median 40 days, range 33 to 46 days).

The mean number of days between botulinum toxin treatment and the six week assessment and between botulinum toxin treatment and the twelve week assessment was 51 days (range 41 to 102 days) and 93 days (range 83 days to 132 days) respectively. The Gantt chart in Figure 4-10 below illustrates the children's time lines of involvement.

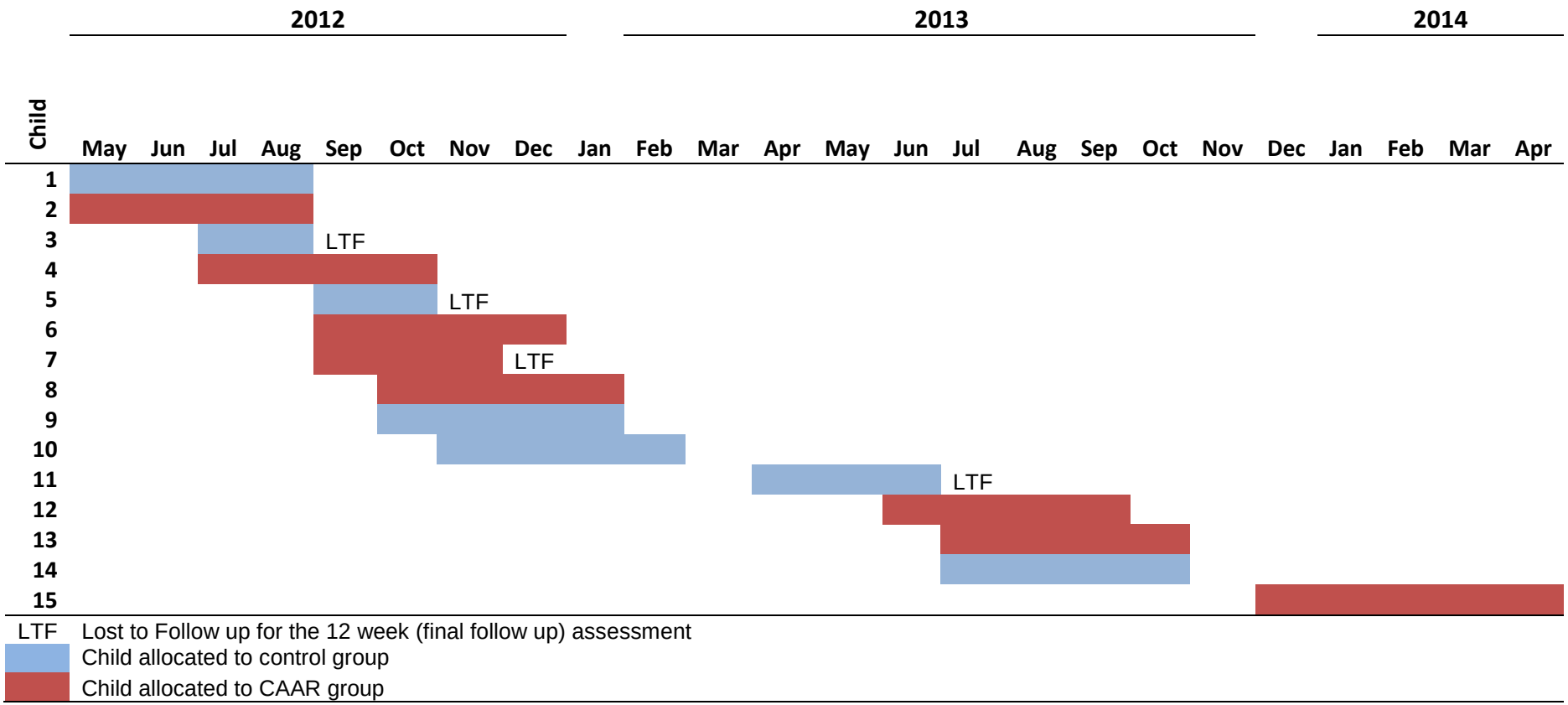
Table 4-4. Record of timings, showing days between study procedures.

Informed consent and baseline assessments all took place on the day of botulinum toxin treatment

			CAAR device		**Visits** (days since botulinum treatment)		
						Assessments	
Child	**Minimised** (days since consent and baseline assessment)	**Allocation**	**Delivery** (days since botulinum toxin treatment)	**Collection** (number of days CAAR in child's home)	**3-week**	**6-week**	**12-week**
1	1	Control	-	-	24	41	91
2	1	CAAR	7	35	29	45	84
3	1	Control	-	-	23	50	LTF
4	0	CAAR	2	41	20	46	86
5	0	Control	-	-	29	41	LTF
6	0	CAAR	2	45	19	48	86
7	0	CAAR	7	40	28	48	LTF
8	0	CAAR	1	46	22	50	98
9	0	Control	-	-	26	42	89
10	0	Control	-	-	17	45	83
11	0	Control	-	-	20	52	LTF
12	0	CAAR	5	33	21	49	83
13	0	CAAR	3	40	21	55	90
14	0	Control	-	-	52	57	97
15	0	CAAR	3	39	*	102	132

LTF: lost to follow up
*unable to establish any contact with family

Figure 4-10. Gantt chart showing children's time lines of involvement (months involved) from informed consent to final assessment at six or twelve weeks.



Up until November 2012, recruitment was progressing at a rate likely to achieve the sample size necessary to power the study adequately (58 participants). From early in 2013 and for the remainder of the year, staffing problems at the LTHT spasticity clinic e.g. maternity leave and retirement of senior consultants reduced the appointments available for children to attend the clinic by up to 67%. Of the fifteen children recruited to the study, one child was recruited at York Hospital NHS Trust (child 3) and two were recruited at Huddersfield and Calderdale NHS Trust (child 6 and child 9).

4.4.3 Participants demographics and clinical details, overall view and within groups comparison

The mean age of the children was 9 years 2 months (standard deviation: 2 years 5 months, range 5 years 2 months to 12 years 9 months). The median age was 9 years 11 months. Children's demographics and clinical details are given in Table 4-5.

Table 4-5. Overall demographics and clinical details of the sample and each group.

	Total (N = 15)	CAAR group (N=8)	Control group (N=7)
Age in years (y) and months (m)			
mean (SD)	9y 2m (2y 5m)	9y 5m (2y 3m)	8y 7m (2y 7m)
median	9y 11m	10y 2m	8y 10m
min	5y 2m	5y 7m	5y 2m
max	12y 9m	12y 9m	11y 9m
Gender			
Male	9	4	5
Female	6	4	2
MACS Levels			
Level II	3	2	1
Level III	5	3	2
Level IV	7	3	4
Distribution			
Bilateral	1	1	0
(R) unilateral	9	6	3
(L) unilateral	5	1	4
Learning impairment			
Present	6	3	3
Not present	9	5	4
Visual impairment			
Present	4	2	2
Not present	11	6	5

The clinical details and demographics of each participant are given in Table 4-6 below.

Table 4-6. Clinical and demographic details of participants.

Participant	Site	Gender	Arms affected	Age years	Age months	MACS level	Home computer games	Allocation	Visual impairment	Learning disability
01	LTHT	Male	(R) unilateral	9	1	MACS IV	Yes	Control	No	Yes
02	LTHT	Male	(R) unilateral	9	11	MACS IV	Yes	CAAR device	Yes	Yes
03	York	Female	(L) unilateral	8	0	MACS IV	Yes	Control	Yes	Yes
04	LTHT	Male	(R) unilateral	9	10	MACS II	Yes	CAAR device	No	No
05	LTHT	Male	(L) unilateral	5	8	MACS III	Yes	Control	No	No
06	H&C	Female	(R) unilateral	12	8	MACS III	Yes	CAAR device	No	No
07	LTHT	Female	(L) unilateral	10	10	MACS III	No	CAAR device	Yes	Yes
08	LTHT	Female	(R) unilateral	7	3	MACS IV	Yes	CAAR device	No	No
09	H&C	Male	(R) unilateral	5	1	MACS IV	No	Control	No	No
10	LTHT	Male	(L) unilateral	11	9	MACS IV	Yes	Control	Yes	Yes
11	LTHT	Female	(R) unilateral	11	5	MACS II	Yes	Control	No	No
12	LTHT	Male	Bilateral	5	7	MACS IV	Yes	CAAR device	No	Yes
13	LTHT	Female	(R) unilateral	11	3	MACS II	No	CAAR device	No	No
14	LTHT	Male	(L) unilateral	8	10	MACS III	Yes	Control	No	No
15	LTHT	Male	(R) unilateral	10	1	MACS III	Yes	CAAR device	No	No

(R): Right;
(L): Left;
MACS: Manual Ability Classification Level;
CAAR: Computer-Assisted Arm Rehabilitation;
LTHT: Leeds Teaching Hospitals NHS Trust;
H&C: Huddersfield and Calderdale.

4.4.4 Minimisation

Table 4-7 shows how each child was allocated in turn to groups through the minimisation process, and shows the summed totals of the stratifying factors accounting for the allocation of each participant. The first child was randomly allocated, because the summed factors were both nil.

The second child's prognostic factors were exactly the same as the first child, so the summed factors were four (for the group into which the first child had been allocated) and nil for the empty group. Therefore the child was minimised into the opposite group of the first child, the group with the lowest summed total of stratifying factors.

The third child had three factors which matched the first two children; this meant the summed factors were three and three, and the child was randomly allocated.

Although the probability of a forced unbalance was 0.2, only once did this occur in fifteen allocations. With summed factors totalling five in the control group against six in the CAAR group, child 13 should have been minimised into the control group but was instead allocated into the CAAR group.

Table 4-7. Participant-by-participant allocation to groups showing the minimisation process.

				Stratification factors				Summed factors	
Participant	Random number	Allocation	Minimisation procedure	Age group	Gender	MACS	Computer games	CAAR	Control
1	1334	Control	Randomised (factors equal)	8 - 10	Male	4	Yes	0	0
2	48	CAAR	Minimised	8 - 10	Male	4	Yes	0	4
3	304	Control	Randomised (factors equal)	8 - 10	Female	4	Yes	3	3
4	934	CAAR	Minimised	8 - 10	Male	2	Yes	3	5
5	1314	Control	Minimised	5 - 7	Male	3	Yes	4	3
6	185	CAAR	Minimised	11 - 12	Female	3	Yes	2	5
7	1421	CAAR	Randomised (factors equal)	8 - 10	Female	3	No	4	4
8	385	CAAR	Minimised	5 - 7	Female	4	Yes	6	7
9	918	Control	Randomised (factors equal)	5 - 7	Male	4	No	5	5
10	938	Control	Randomised (factors equal)	11 - 12	Male	4	Yes	9	9
11	973	Control	Minimised	11 - 12	Female	2	Yes	9	6
12	34	CAAR	Minimised	5 - 7	Male	2	Yes	8	12
13	1696	CAAR	Forced unbalance	11 - 12	Female	2	No	6	5
14	151	Control	Minimised	8 - 10	Male	3	Yes	13	12
15	141	CAAR	Randomised (factors equal)	8 - 10	Male	2	Yes	15	15

4.4.5 Success of blinding procedure

The allocation of children to groups was revealed to the blinded assessor on two occasions, firstly by child 4 and secondly by the parent of child 7. Of the other thirteen children, the blinded assessor did not know the allocation at any time but correctly guessed the allocation of six children (46%). A binomial test of proportions with significance set at 0.05 produces a confidence interval of 34% and 58%, suggesting that blinding was successful.

If the two children that revealed their allocation were included in the binomial test of proportions, the blinded assessor correctly identified the allocation of eight children (53%). The confidence interval was 42% to 64%, suggesting that successful blinding of assessments was achieved even with the two revealed allocations.

4.4.6 Use of CAAR games device: amount of time played

Figure 4-11 illustrates the total number of minutes that the device was used by each child.

Figure 4-11. Bar chart illustrating the number of minutes that the CAAR device was used by each child during the home deployment.

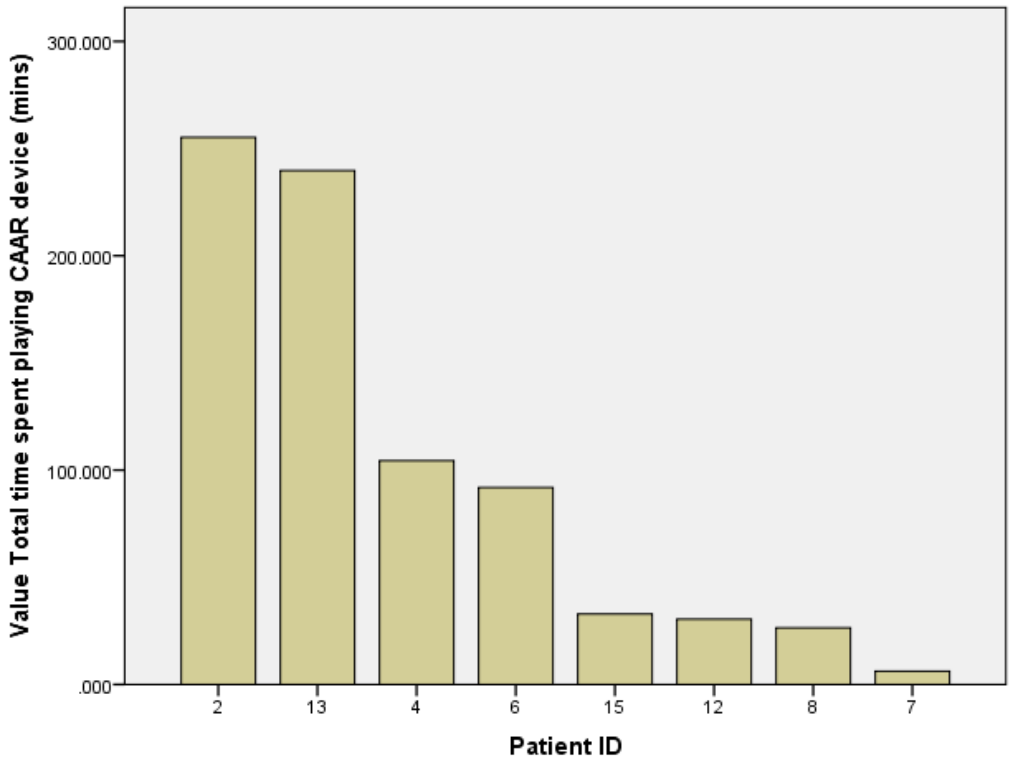


Table 4-8 below shows the number of days and the amount of time each child used the CAAR device. The mean total use between the children was 99 minutes. The mean daily amount of time the device was played was seven minutes, substantially less than the 30

minutes per day that was suggested to parents. Child 2 and child 13 used the games device the greatest amount, with a total of over four hours each (daily mean 10min 40s and 8min 20s, playing on 24 and 29 days respectively), while child 7 played the games the least (for a total of six minutes, taking place on four days in the first week only). The highest daily mean was child 15, who played the games for almost 11 minutes a day, but the child played the games on only three out of 40 days on which he had the opportunity.

The mean number of days the device was in the house was 40 days, and the device was played on a mean of 14 days. Half the children used the device for three or fewer of the six weeks, with one child using the device in the first week only, for a total of six minutes.

Table 4-8. Amount of use by the participants of the CAAR games device in minutes and number of days used.

Child	Weekly use of device (minutes)							Total use (minutes)	Mean number of minutes played per day	Days played
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7			
2	57.8	48.8	43.0	47.3	59.2	-	-	256.2	10m 40s	24
4	32.0	40.2	14.8	10.2	7.1	1.0	-	105.4	5m 51s	18
6	41.9	22.9	5.7	5.7	1.5	4.8	10.5	92.9	4m 2s	23
7	6.0	-	-	-	-	-	-	6.0	1m 30s	4
8	13.4	10.5	2.6	-	-	-	-	26.5	3m 47s	7
12	27.1	3.4	-	-	-	-	-	30.5	6m 6s	5
13	47.8	28.7	57.2	61.9	0.0	45.9	-	241.5	8m 20s	29
15	14.1	-	18.8	-	-	-	-	32.9	10m 58s	3

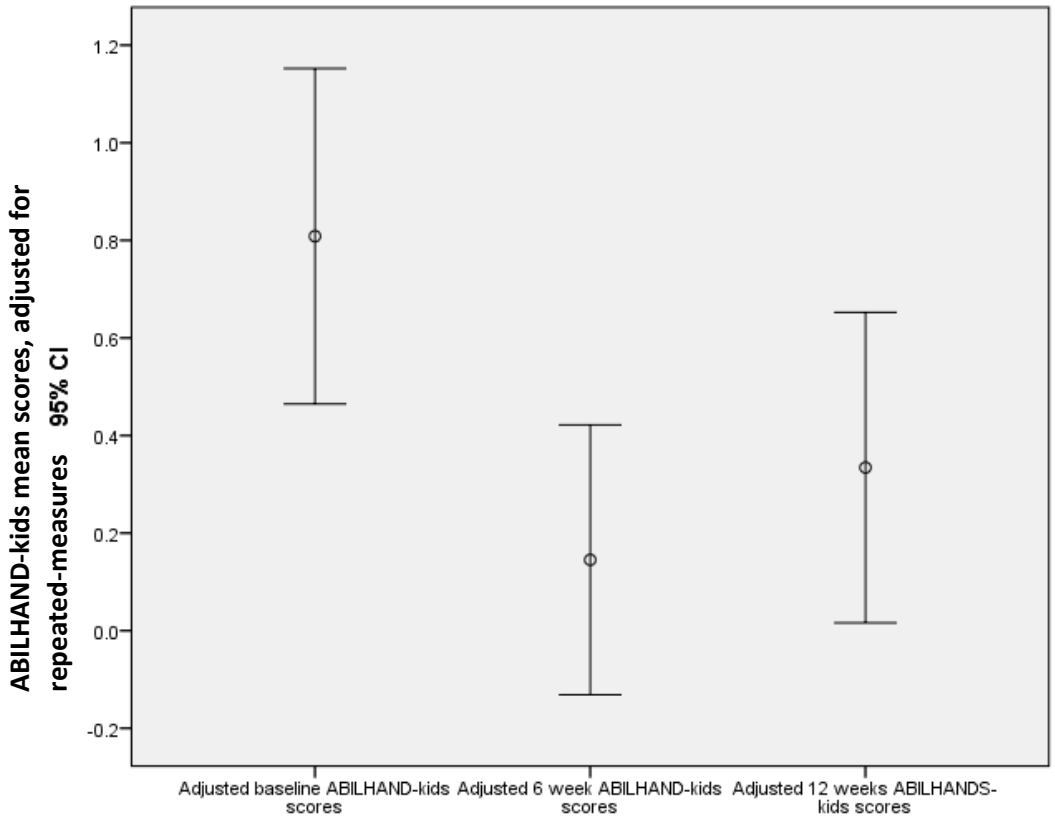
m: minutes
s: seconds

4.4.7 Outcome measures: scores, comparisons and tests of statistical analyses

4.4.7.1 Primary outcome measure: the ABILHAND-kids

Figure 4-12 below shows an error bar graph illustrating the within-participants differences across time points for the mean scores in ABILHAND-kids scores, adjusted for repeated measures to remove the unsystematic variation from the data plotted (Field, 2009). The error bars show the 95% confidence interval, for which there is no overlap between the baseline and six week time point.

Figure 4-12. Error bar graph illustrating repeated-measures (across time points) adjusted ABILHANDS-kids mean scores for all participants.



The ABILHAND-kids results for all participants after transformation into linear scores are given in Table 4-9 below. The table gives each participant's individual scores, the mean scores for all participants (as illustrated in Figure 4-12 above) and medians. No minimum clinically significant difference is given for the ABILHAND-kids so a change greater than the standard error (mean SE = 0.44) was accepted as clinically significant.

A clinically significant improvement was observed in two children, one from the control group and one from the CAAR group. However, nine of the children showed a clinically significant deterioration in activity performance, five in the control group and four in the CAAR group. Twelve children therefore did not show any improvement in activity limitation following botulinum toxin treatment.

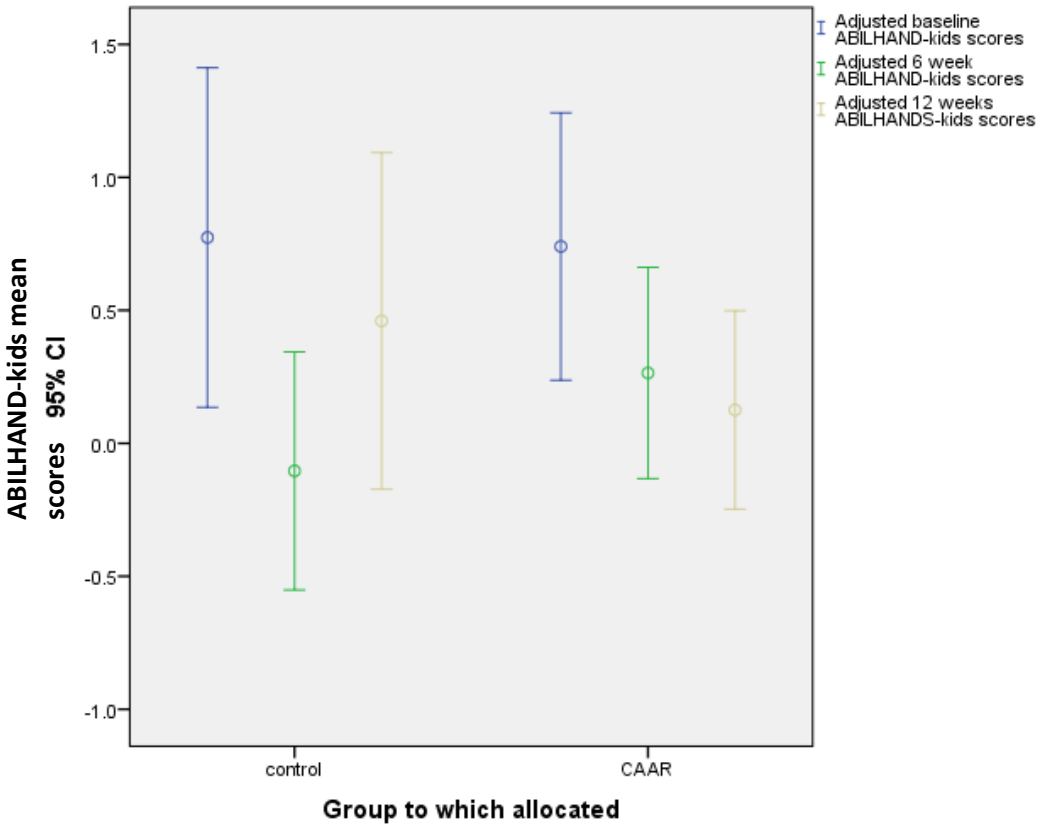
Table 4-9. Change scores for the primary outcome measure (the ABILHAND-kids).

			ABILHAND-kids change in outcome score		
Child	**Clinical change**	**Allocation**	**Baseline**	**6-week**	**12-week**
1	Improved	Control	0.592	1.203	1.065
2	Deteriorated	CAAR	0.509	0.004	-0.453
3	Deteriorated	Control	-0.843	-3.184	LTF*
4	No change	CAAR	0.645	0.852	0.509
5	No change	Control	-0.653	-0.671	LTF*
6	Deteriorated	CAAR	1.172	-1.206	-0.164
7	Improved	CAAR	-0.332	0.352	LTF*
8	No change	CAAR	2.025	1.963	0.852
9	Deteriorated	Control	0.32	-0.501	-0.332
10	Deteriorated	Control	1.763	1.026	3.183
11	Deteriorated	Control	2.365	0.852	LTF*
12	Deteriorated	CAAR	-1.018	-2.393	-3.184
13	Deteriorated	CAAR	3.183	2.634	3.183
14	Deteriorated	Control	1.718	0.394	2.154
15	No change	CAAR	0.68	0.852	0.852
		Mean (SD)	0.8084 (1.23)	0.145 (1.56)	0.334 (1.90)
		Median	0.645	0.394	0.509

*LTF: Lost to Follow Up. Score for 6 weeks inserted for analyses because those lost to follow up were assumed to have made no change.

A comparison of the mean scores for control and CAAR groups is illustrated in Figure 4-13 below. The overlap of the 95% confidence error bars suggest that botulinum treatment had no effect on arm activity limitation of either group across time points or between groups at any time point.

Figure 4-13. Error bar graph illustrating adjusted repeated-measures (across time points) ABILHANDS-kids mean scores for control and CAAR groups.



The repeated-measures ABILHAND-kids scores at each time point for each group (control and CAAR group) are shown in Table 4-10.

Table 4-10. Descriptive statistics for the control and CAAR groups ABILHAND-KIDS scores across time points.

Allocation	Statistic	Baseline	Six weeks	Twelve weeks
Control	**Mean (SE)**	0.75 (0.47)	-0.13 (0.58)	0.44 (0.79)
	Median	0.59	0.39	0.85
	Minimum	-0.84	-3.18	-3.18
	Maximum	2.37	1.20	3.18
	Range	3.21	4.39	6.37
CAAR	**Mean (SE)**	0.86 (0.46)	0.38 (0.57)	0.24 (0.62)
	Median	0.66	0.60	0.43
	Minimum	-1.02	-2.39	-3.18
	Maximum	3.18	2.63	3.18
	Range	4.20	5.03	6.37

SE: standard error

An independent t-test showed that at baseline there was no difference between the control group and the CAAR group ABILHAND-kids scores (mean*(SE)* = 0.75(0.47) and 0.86(0.46) respectively ($t(13) = -0.160$, $p=0.875$. Scores decreased in both groups at six weeks. At twelve weeks the CAAR group scores had deteriorated further, but the control group scores had improved, although not to the baseline level.

The mean scores for the control group across time points at baseline, six weeks and twelve weeks were 0.75, -0.13 and 0.44 respectively. The scores for the CAAR group were 0.86, 0.38 and 0.24.

The differences between groups and time points were analysed using the statistical tests described in Statistical analyses on page 227.

4.4.7.1.1 Comparing group means: mixed design ANCOVA

Before performing statistical tests on the results to determine the statistical significance of the ABILHAND-kids results, it is essential to test for assumptions on which parametric statistical calculations are based. First of all, the ABILHAND-kids baseline scores were assessed for a Normal distribution.

4.4.7.1.1.1 Testing assumptions: Normal distribution

The descriptive statistics in Table 4-11 show a median that differed markedly from the mean score, suggesting a non-Normal distribution. However, converting the scores to z-scores shows that all scores fall within the 95% confidence interval (two standard deviations), suggesting a Normal distribution. Skewness gives a guide to whether Normal distribution exists by converting to z-scores (skewness statistic divided by Standard Error) and looking for a value of between -1.96 to 1.96 (Field, 2009). In this case, the value is 0.40, suggesting a Normal distribution.

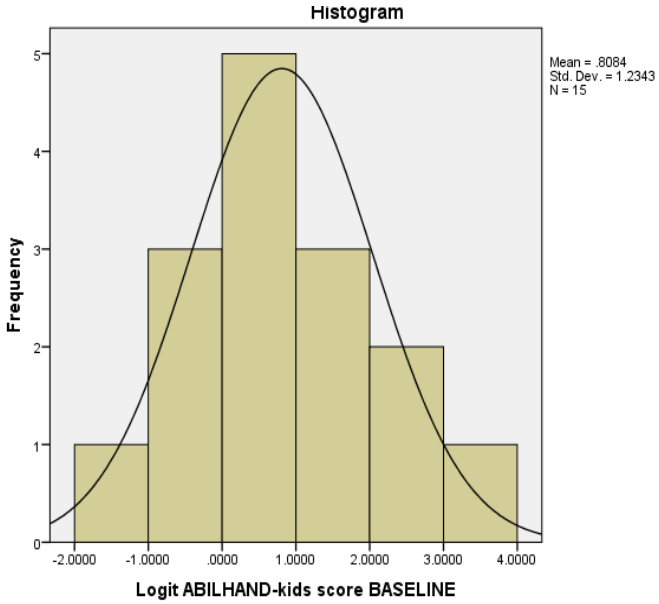
Table 4-11. Descriptive statistics for ABILHAND-kids baseline scores

	Statistic	**Standard Error**
Mean (SD)	0.81 (1.23)	0.32
Median	0.65	
Skewness	0.23	0.58
Kurtosis	-0.63	1.12

SD: standard deviation

Figure 4-14 illustrates the histogram and Normal curve for the ABILHAND-kids baseline scores. This suggests a slight positive skew, and a possible non-Normal distribution.

Figure 4-14. Histogram and Normality curve for baseline ABILHAND-kids scores.



Finally, Kolmogorov-Smirnov and Shapiro-Wilk tests of Normality were conducted on the data (Field, 2009). The results of these tests are given in Table 4-12 below. The non-significant p-value for each test suggests a distribution that does not deviate significantly from a Normal distribution.

Table 4-12. Results of Kolmogorov-Smirnov and Shapiro-Wilk tests of Normality on baseline ABILHAND-kids scores.

Normality test	Statistic	df	p-value
Kolmogorov-Smirnov	0.141	15	0.200
Shapiro-Wilk	0.965	15	0.779

df: degrees of freedom

Results at six weeks and twelve weeks were also evaluated and found to have acceptable Normality. The histogram for six weeks illustrates a potential negative skew (see Figure 4-15), but other tests indicated that there was no deviation from a Normal distribution: *z*-scores for the six and twelve weeks' skewness were -1.24 and -0.81 respectively, and the results of Kolmogorov-Smirnov and Shapiro-Wilk tests for Normality given in Table 4-13 below are non-significant.

Figure 4-15. Histogram for six week ABILHAND-kids scores showing a potential negative skew.

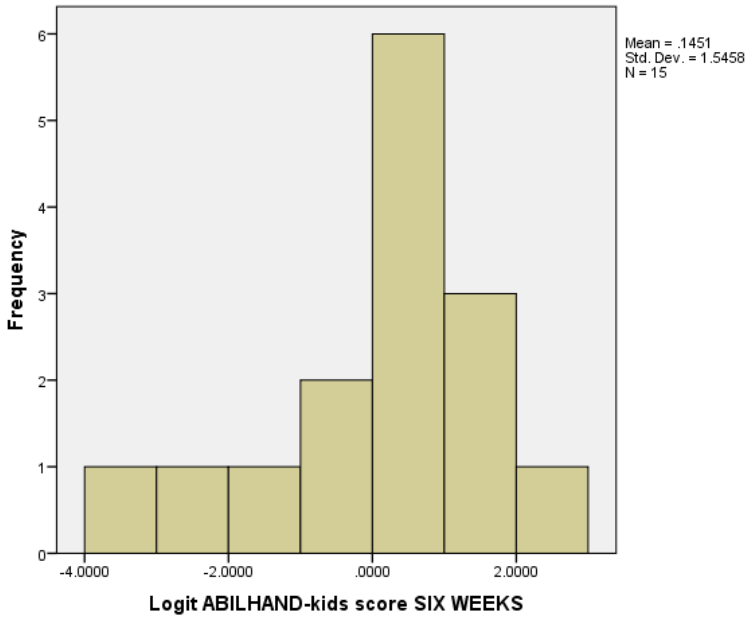


Table 4-13. Tests for a Normal distribution of six and twelve week ABILHAND-kids results.

	Tests of Normality					
	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	p-value	Statistic	df	p-value
Six weeks	0.153	15	0.200	0.95	15	0.517
Twelve weeks	0.16	15	0.200	0.92	15	0.195

df: degrees of freedom

These explorations of the ABILHAND-kids outcome scores suggest that a parametric statistical approach is acceptable. Therefore, a one-way mixed design ANCOVA was conducted to evaluate the effects of botulinum toxin treatment within groups across time points and to evaluate whether any differences between groups occurred due to playing the games system. The covariates of age, gender, MACS levels and use of commercial games systems were included in the ANCOVA.

4.4.7.1.1.2 Further assumption testing

The SPSS output from a mixed design ANCOVA includes the results of three additional tests for assumptions that must be met for the validity of the ANCOVA. For the between-groups tests, Levene's test of Equality of Error Variances evaluates homogeneity of variances at each time point. These were acceptable (p=0.704, p=0.645, p=0.597 respectively) and Box's Test of Equality of Covariance Matrices tests the homogeneity of the covariance matrices which underpin the mathematical workings of the ANCOVA. This was also non-significant

(*p*=0.983). Finally, the repeated measures test assumes homogeneity of variance (sphericity) which was tested with Mauchly's Test of Sphericity. The result of this test was non-significant (*p*=0.163), indicating that sphericity was acceptable.

4.4.7.1.2 ANCOVA results

The results of the repeated-measures (within-participants) ANCOVA are given in Table 4-14 below. These indicate that the differences in ABILHAND-kids scores between time points shown in Table 4-9 are non-significant F(2,18) = 0.807, *p*=0.462, adjusting for the covariates of gender, age, arm disability (MACS) and use of home computer games which were all included in the ANCOVA. This suggests that botulinum treatment made no difference to the children's upper limb activity limitation up to the twelve week assessment. The interaction effect between time points and allocation was non-significant F(2,18)=0.138, *p*=0.699, indicating that use of the CAAR games had no influence on arm activity limitation.

Table 4-14. Results of within-participants (repeated-measures) ANCOVA showing significance of comparisons between time points.

Sum of Squares	df	Mean Square	F	*p*-value
Time points				
0.609	2	0.305	0.807	0.462
Time points against allocation				
0.276	2	0.138	0.365	0.699
Time points against age				
3.027	2	1.514	4.01	0.036
Error (time points)				
6.794	18	0.377		

The ANCOVA results indicate that there was no interaction between any covariate across time points except for an interaction between age and time points (F(2,18)=4.01, *p*=0.036. Planned contrasts examining the effects of age on changes in arm activity limitation between baseline and six weeks, and baseline and twelve weeks, were both non-significant (F(1,9)=0.646, *p*=0.442; F(1,9)=2.619, *p*=0.14 respectively).

The result of the between-participants ANCOVA is given in Table 4-15 below. This reveals that there was no difference between groups at each time point F(1,8)=0.011, *p*=0.919. This indicates that use of the CAAR gaming device made no impact on the CAAR group's activity limitation.

Table 4-15. Results of between-groups (repeated-measures) ANCOVA.

	Sum of Squares	df	Mean Square	F	*p*-value	Partial Eta Squared
Allocation (CAAR or control)	0.027	1	0.027	0.011	0.919	0.001
Error	19.377	8	2.422			

df: degrees of freedom
F: F statistic

Finally, the ANCOVA Tests Of Between-Subjects Effects show that the covariates of age, gender, MACS levels and use of commercial home computer games were not significantly related to differences between groups outcome scores.

4.4.7.2 Secondary outcome activity limitation measures

4.4.7.2.1 The Canadian Occupational Performance Measure (COPM)

4.4.7.2.1.1 COPM activities and scores for each participant

Table 4-16 below shows the outcome goals that were developed in semi-structured interviews with the children's parents, as standard with COPM use. The goals are activities involving the upper limb at which children showed limitations. Included alongside each of the activities are the scores out of ten at which the child's performance (and satisfaction with that performance) was graded by the parent(s) at all three time points (baseline, six weeks and twelve weeks).

Table 4-16. Upper limb goals selected by parents and scored out of ten for performance and satisfaction at baseline, six weeks and twelve weeks.

		Performance scores (out of ten)			Satisfaction scores (out of ten)		
Child	**Goals selected by parents at baseline**	**Baseline**	**6 week**	**12 week**	**Baseline**	**6 week**	**12 week**
1	Catching a football	5	5	7	8	5	7
1	Opening a car door	3	2	2	4	3	3
1	Putting trousers on	5	4	5	5	4	6
1	Putting a vest on	7	6	8	9	7	9
1	Putting shoes on	3	2	1	3	3	1
2	Putting on socks	3	6	4	5	7	4
2	Knife and fork use: cutting up a sausage	2	6	6	3	8	8
2	Fastening trousers: school trousers, with a hook	4	7	6	4	9	7
2	Using a ruler to underline a few words	2	4	3	3	5	3
3	Able to pull trousers up	4	5	LTF	9	7	LTF
3	Independently put on school uniform	2	5	LTF	7	7	LTF
3	Spontaneous use of impaired arm	1	2	LTF	8	3	LTF
4	Riding bike	3	3	3	8	2	7
4	Putting on a short sleeve shirt	3	9	2	6	9	6
4	Donning socks	2	4	7	4	3	9
4	Using cutlery with both hands	1	1	2	2	0	5
4	Spontaneous use of impaired arm	3	6	7		6	8
5	Taking own T-Shirt off	4	5	LTF	2	5	LTF
5	Spontaneous use of impaired arm	2	8	LTF	2	5	LTF
5	Holding paper still when writing with non-impaired arm	7	8	LTF	5	8	LTF
5	Holding computer game operating handles when playing computer game	4	8	LTF	4	8	LTF

5	Catching a ball	4	5	LTF	3	5	LTF
6	Washing own hair	5	6	6	7	4	6
6	Fastening top button on school shirt	1	1	3	1	1	1
6	Spreading butter on bread	4	1	3	4	1	1
6	Cracking an egg	1	1	1	3	1	1
6	Spontaneous use of impaired arm	1	1	3	4	1	1
7	Brushing own hair independently	5	3	LTF	7	10	LTF
7	Donning own ankle socks	8	7	LTF	10	10	LTF
7	Spontaneous use of impaired arm	5	6	LTF	5	10	LTF
7	Use of a knife (with a fork) when eating meals	3	4	LTF	2	7	LTF
7	Breast stroke (arm action) when swimming	7	3	LTF	10	10	LTF
8	Independently putting on socks	5	4	5	4	5	5
8	Use of a knife (with a fork) when eating meals	3	5	3	2	3	3
8	Spontaneous use of impaired arm	6	5	3	7	6	3
8	Holding paper still when writing with non-impaired arm	6	5	3	5	5	3
9	Spontaneous use of impaired arm	2	1	2	1	1	3
9	Independently going to the toilet (sitting down), not including washing hands	2	3	2	1	1	3
9	Performing a grip hold in judo	1	1	2	1	1	3
9	Catching a tennis ball two-handed	3	2	4	3	2	3
9	Eating with a knife and fork	1	1	2	1	1	3
10	Buttoning two buttons of a shirt	6	9	9	5	10	9
10	Hooking a trouser fastening	2	6	9	3	10	9
10	Spontaneous use of impaired arm	4	7	7	2	9	8
10	Use of knife (with fork) to cut soft food	2	3	2	1	8	5

10	Bouncing ball with alternate hands	6	9	9	6	10	9
11	Independent fastening of zips	1	1	LTF	1	1	LTF
11	Use of ruler/protractor and other stationery	1	4	LTF	1	4	LTF
11	Knife and fork use	4	6	LTF	6	6	LTF
11	Independently washing hair	5	3	LTF	5	2	LTF
11	Spontaneous use of impaired arm	1	5	LTF	1	3	LTF
12	Cleaning teeth	3	5	10	3	10	10
12	Combing hair	5	10	10	8	10	10
12	Using an iPad	7	10	10	10	10	10
13	Crafts - cutting paper	6	8	7	8	9	6
13	Cutting chicken with a knife and fork	2	5	5	3	5	4
13	Putting on socks	5	7	8	5	8	7
13	Spontaneous use of impaired arm	6	7	4	7	8	2
13	Holding a tray or photo	7	7	7	6	8	7
14	Cutting meat with knife and fork	4	5	7	6	6	9
14	Spontaneous use of impaired arm	2	4	4	2	4	4
14	Buttoning up the buttons on a polo shirt	5	2	4	3	3	4
14	Crafts: cutting shapes	6	2		4	4	
14	Doing up the button on his jeans	3	2	7	3	3	7
15	Fastening buttons	4	4	6	3	6	6
15	Using arm for support to get in and out of the bath	3	4	5	3	6	5
15	Putting on right sock	3	4	6	3	7	6
15	Shampooing hair	3	4	6	3	7	6
15	Stabilising plates and paper when eating or writing	3	4	6	3	6	6

COPM outcome scores for each participant are calculated as a mean of scores for all goals selected by the parents for each participant. The change in scores from baseline to each of the follow up assessments is shown in Table 4-17. A change score of 2.0 or greater represents a clinically-significant change, so these results reveal that four children out of the fifteen showed a clinically significant improvement at six weeks, two in the CAAR group and two in the control group. Two of these four children had maintained the improvement at twelve weeks, and a fifth child had achieved a clinically significant improvement by twelve weeks.

Table 4-17. Changes in parent's perception of the child's performance and changes in parent's satisfaction.

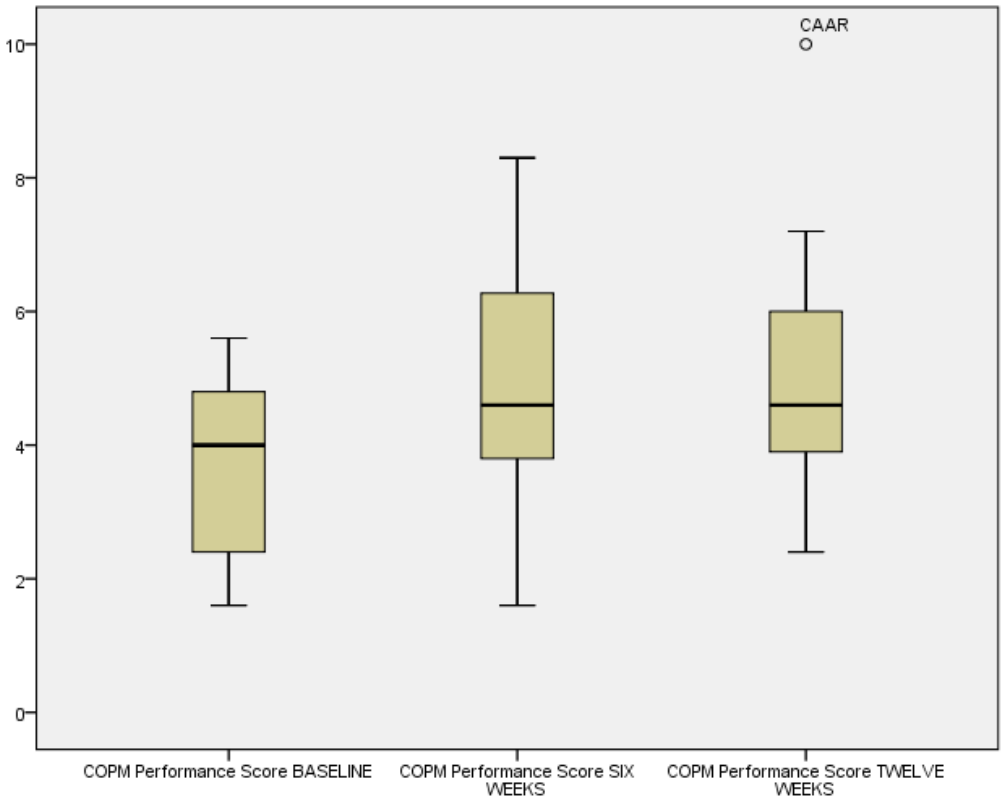
		Change score: performance		Change score: satisfaction	
Child	**Allocation**	**6 weeks**	**12 weeks**	**6 weeks**	**12 weeks**
01	Control	-0.8	0	-0.4	0.4
02*[T]	CAAR	3	2	3.5	1.75
03	Control	1.67	LTF	-2.34	LTF
04*	CAAR	2.2	1.8	0	2
05*	Control	2.6	LTF	3	LTF
06	CAAR	-0.4	0.8	-2.2	-1.8
07	CAAR	-1	LTF	2.6	LTF
08	CAAR	-0.25	-1.5	0.25	-1
09	Control	0	0.8	-0.2	1.6
10*[T]	Control	2.8	3.2	6	4.6
11	Control	1.4	LTF	0.4	LTF
12	CAAR	0	1.2	3	3
13	CAAR	1.4	-0.15	1.8	-0.4
14	Control	-1	1.25	1.4	2.65
15[T]	CAAR	0.8	2.6	3.4	2.8

* clinically significant improvement at 6 weeks

[T] clinically significant improvement at 12 weeks

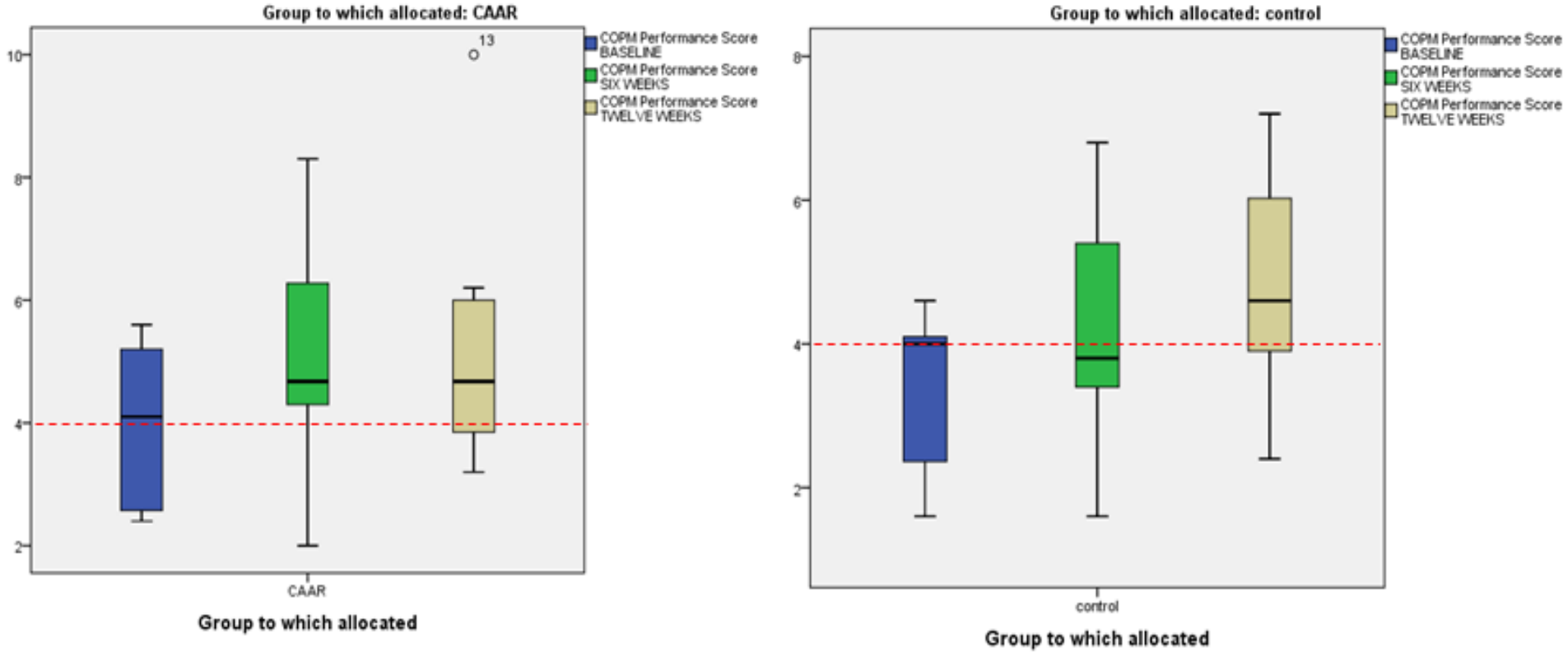
The box plot in Figure 4-17 below illustrates within-participants changes across time points.

Figure 4-17. Box plots illustrating within-participants changes in COPM scores across time points.



A box plot comparison of between-groups and across time points (within-groups) COPM scores is illustrated in Figure 4-18 below.

Figure 4-18. An SPSS box plot between-groups and within-groups (across time points) comparison of COPM scores. Note scales for each box plot are not aligned (reference is drawn at 4.0 to aid comparisons).



4.4.7.2.1.2 Within-groups comparisons of COPM scores (between time points)

The COPM's ordinal level outcome scores do not meet the criteria for parametric statistical calculations therefore non-parametric ANOVAs were used to determine whether there were any differences within groups between time points. The assumptions for non-parametric testing are that samples are randomly selected and that no participants are in more than one category or group. The exception to the latter assumption is when participants are subjected to repeated measures tests (Pallant, 2007).

The Friedman Test is a non-parametric one-way repeated-measures ANOVA (Pallant, 2007). The results of this statistical test showed that the difference across time points shown by all participants (4.0 at baseline, 4.6 at six weeks and 4.6 at twelve weeks, as illustrated in Figure 4-17) reveal that a significant difference exists for COPM scores between time points, $\chi^2(2, 15) = 6.778$, p=0.031.

An ANOVA indicates whether a difference is significant between two or more of several groups, but does not indicate the whereabouts or the direction of the difference. This is the purpose of *post hoc* testing.

The median scores suggest that the significant finding applies between the baseline score (median = 4.0) and six week scores (median = 4.6), and between the baseline and twelve week score (median = 4.6). To evaluate this, a non-parametric within-groups Wilcoxon Signed Ranks Test was performed.

4.4.7.2.1.2.1 Significance testing, post hoc tests: the non-parametric Wilcoxon Signed Ranks Test repeated measures test

Because a number of repeated measures are being performed (baseline to six weeks, and baseline to twelve weeks), it is necessary to adjust the *p*-value to prevent a Type I error. This adjustment is the Bonferroni correction (Bland and Altman, 1995) and is calculated by dividing the required significance level by the number of comparisons (Field, 2009) i.e. a significance level of 0.05/2, or 0.025. The one-tailed significance level is appropriate because it was identified *a priori* that the outcome is investigating the hypothesis of improved activity limitation following botulinum toxin treatment (Field, 2009).

The effect size (r) is calculated by dividing the z value by the square root of the number of observations i.e. the number of participants observed over two time points (z-value divided by 15 x 2) (Field, 2009, Pallant, 2007). An effect size of 0.3 to 0.5 is a small effect (Field, 2009).

The results of the Wilcoxon Signed Ranks Test revealed a statistically-significant increase in COPM scores at the six week time point (median = 4.6) compared to baseline (median = 4.0, z=-2.199, p=0.013). This p-value is still lower than the Bonferroni-adjusted level of significance of 0.025. The effect size is small (r = 0.4) (Field, 2009). The change in scores is less than 2 and therefore is not clinically significant. At twelve weeks (median = 4.6), the scores were still significantly above the baseline scores z=-2.608, p=0.003, effect size was still small (r = 0.48). The difference is not clinically significant, with a change score of less than 2.0 from the baseline score.

4.4.7.2.1.3 Between-groups comparisons of COPM scores

The median scores for each group at each time point are given in Table 4-18.

Table 4-18. Median COPM scores for control and CAAR groups at each time point.

	Baseline		Six week		Twelve week	
	Control	**CAAR**	**Control**	**CAAR**	**Control**	**CAAR**
Median	4.0	4.1	3.8	4.7	4.6	4.7

The Kruskal-Wallis Test is a non-parametric between-groups ANOVA (Field, 2009, Pallant, 2007). This was used to test for the differences between group medians at each time point, as illustrated in Figure 4-18 above. The Kruskal-Wallis Test revealed that there was no difference in COPM scores between groups at each time point (see Table 4-19), with all comparisons non-significant at the 0.05 level: difference at baseline = 0.1 (χ^2(1, n=15)=1.638, p=0.201), at six weeks= 0.9 (χ^2(1, n=15)=1.495, p=0.221), and at twelve weeks = 0.1 (χ^2(1, n=15)=0.03, p=0.862).

Table 4-19. Results of between-groups ANOVA (Kruskal-Wallis Test) for differences in COPM scores (activity levels) between groups at each time point.

	Baseline	Six Weeks	Twelve Weeks
Chi-Square	1.638	1.495	0.03
df	1	1	1
***p*-value**	0.201	0.221	0.862

df: degrees of freedom

4.4.7.3 Kinematic analyses

Apart from the children lost to follow up, CPKAT results are missing at twelve weeks for child 15 and for all time points for child 12. Updates to the University of Leeds' computer operating systems caused a malfunction of the CPKAT laptop for the twelve week assessment of child 15, preventing kinematic assessment of this participant. Child 12 was a wheelchair user whose wheelchair and home seating impeded optimal positioning of the joystick, preventing CPKAT assessment. Their missing scores were substituted for the six week scores, as it was assumed that no change took place.

4.4.7.3.1 Pentagram results

Median scores of each participant for all Pentagram kinematic parameters are shown in Table 4-20. There do not appear to be significant differences between scores in any of these parameters from one time point to the next.

Table 4-20. Raw scores of Pentagram kinematic assessments for all participants at each time point.

	Movement time (s)			Path length (mm)			Path length time (s)			NJ (smoothness, no units)		
Child	**Baseline**	**6 weeks**	**12 weeks***	**Baseline**	**6 weeks**	**12 weeks***	**Baseline**	**6 weeks**	**12 weeks***	**Baseline**	**6 weeks**	**12 weeks***
1	2.56	2.71	3.07	192.64	224.88	305.86	2.96	3.69	3.615	1803.84	2838.90	2168.66
2	2.565	2.43	2.77	188.97	291.91	205.63	2.885	3.01	3.03	1374.91	1282.39	1860.38
3	3.425	3.13	0	152.46	144.14	0	3.58	3.49	0	2640.67	2840.11	0
4	2.125	2.02	1.44	167.94	210.08	158.88	2.61	2.49	1.95	1337.25	1275.90	920.32
5	2.795	2.41	0	140.74	152.15	0	2.885	2.98	0	1904.75	1984.13	0
6	1.58	1.51	1.4	145.01	138.70	135.26	2.08	2.15	1.84	808.57	926.80	855.76
7	6.06	7.09	0	310.48	269.87	0	6.34	8.25	0	4813.51	17495.21	0
8	2.66	5.535	3.1	171.02	287.11	186.59	2.92	5.595	3.47	1563.12	6405.06	2152.62
9	3	2.21	2.745	173.88	189.89	263.17	3.23	2.5	3.685	1939.06	1273.68	3692.34
10	2.45	3.07	2.96	170.91	231.92	226.06	2.77	3.615	3.15	1625.37	2449.87	2151.11
11	1.7	1.77	0	134.99	154.50	0	1.75	2.195	0	507.14	958.03	0
13	1	1.53	1.625	132.85	152.92	144.87	1.63	1.92	2	462.60	572.60	799.98
14	1.705	1.8	1.645	143.60	172.12	150.77	2.42	2.175	1.83	1139.90	1292.17	884.68
15	1.79	1.27	0	138.17	134.18	0	1.91	1.7	0	731.30	682.33	0
Median	**2.505**	**2.31**	**2.5775***	**160.20**	**181.00**	**156.70***	**2.83**	**2.74**	**3.01***	**1469.02**	**1287.28**	**1922.26***

s: seconds; mm: millimetres; NJ: Normalised Jerk

* missing values at 12 weeks were children Lost To Follow-up. Six week results were substituted for statistical evaluations because it was assumed that children showed no improvement across this time point.

4.4.7.3.2 Pentagram analyses

Participants lost to follow up at twelve weeks were assumed to show no change between consecutive time points, so children's six-week scores were substituted for the missing values at twelve-week time point.

Testing for a Normal distribution, the Kologorov-Smirnov Test and the Shapiro-Wilk Test were significant on both the Movement Time parameter and the Path Length parameter, indicating deviation from a Normal distribution. The skewness *z*-value was 4.88, well outside the 1.96 to -1.96 range that indicates a Normal distribution. An examination of the histograms also strongly indicated a skewed distribution.

The data were log transformed (Field, 2009) and histograms generated for each variable. All histograms showed acceptable distribution for parametric statistical operations, and a mixed design ANCOVA was therefore performed including the covariates of gender, age, arm disability (MACS) and use of home computer games.

4.4.7.3.3 Mixed design ANCOVA results for Pentagram scores

The results of the repeated-measures (within-participants) ANCOVA for the Pentagram parameters are given in given in Table 4-14Table 4-21. The results of the ANOVA show that there were no differences across time points for any parameter, adjusting for the covariates of age, gender, MACS and the use of commercial computer games. For example, there was no difference Movement Time (MT) ($F(2,16) = 1.664$, $p=0.224$). These results suggest that botulinum treatment made no difference to the children's upper limb kinematics for any of the Pentagram parameters across time points. There was no significant interaction effect between allocation and time points e.g. for Movement Time $F(2,16)=0.472$, $p=0.632$, suggesting that use of the CAAR games had no influence on the children's upper limb kinematics as measured by the Pentagram task.

Table 4-21. Results of within-participants (repeated-measures) ANCOVA for Pentagram scores showing significance of comparisons between time points.

	Sum of Squares	df	Mean Square	F	*p*-value
Movement Time					
Time points	0.018	2	0.009	1.664	0.224
Time points against allocation	0.005	2	0.003	0.472	0.632
Error (time points)	0.086	16	0.005		
Path Length					
Time points	0.002	2	0.001	0.381	0.690
Time points against allocation	0.016	2	0.008	2.695	0.098
Error (time points)	0.046	16	0.003		
Path Length Time					
Time points	0.003	2	0.002	0.083	0.921
Time points against allocation	0.045	2	0.022	1.066	0.367
Error (time points)	0.336	16	0.021		
Normalised Jerk (smoothness)					
Time points	0.189	2	0.094	1.277	0.306
Time points against allocation	0.374	2	0.187	2.533	0.111
Error (time points)	1.183	16	0.074		

df: degrees of freedom
F: F statistic

The result of the between-groups ANCOVA is given in Table 4-22 below. There was no difference between groups for any of the Pentagram parameters e.g. Movement Time $F(1,8)=0.010$, $p=0.922$. This shows that use of the CAAR gaming device made no impact on the CAAR group's upper limb kinematics.

Table 4-22. Results of between-groups ANCOVA on Pentagram scores.

	Sum of Squares	df	Mean Square	F	*p*-value
Movement time					
Allocation (CAAR or control)	0.027	1	0.027	0.011	0.919
Error	19.377	8	2.422		
Path Length					
Allocation (CAAR or control)	0.037	1	0.037	1.884	0.207
Error	0.159	8	0.020		
Path Length Time					
Allocation (CAAR or control)	0.003	1	0.003	0.76	0.790
Error	0.356	8	0.044		
Normalised Jerk (smoothness)					
Allocation (CAAR or control)	0.048	1	0.048	1.629	0.238
Error	0.236	8	0.029		

df: degrees of freedom
F: F statistic

4.4.7.3.4 CPKAT Tracing task raw scores for each participant

The Tracing task raw scores for each participant at each time point are given in Table 4-23. Twelve-week scores for children lost to follow-up (at the 12 week assessment) have had the six week scores inserted because it was assumed that no change had taken place.

Table 4-23. Raw scores of Tracing task kinematic assessment for all participants at each time point.

	Path Length			Path Length Time			NJ			Accuracy			TPA		
Child	Baseline	6-weeks	12-week	Baseline	6-weeks	12-week	Baseline	6-weeks	12-week	Baseline	6-weeks	12-week	Baseline	6-weeks	12-week
1	797.47	747.42	841.60	31.97	23.65	21.73	212023.18	110108.10	82046.60	3.77	2.95	5.31	120.43	96.33	145.75
2	393.25	511.05	588.12	14.23	8.38	11.35	39082.34	10520.64	24559.68	3.13	5.48	5.06	45.35	55.76	81.96
3	459.01	547.85	547.85	25.89	26.84	26.84	124309.94	157775.11	157775.11	2.85	2.64	2.64	81.41	83.50	83.50
4	481.50	652.66	659.63	19.68	35.55	30.09	77042.39	240179.59	182434.64	2.48	2.49	2.38	48.89	90.39	84.48
5	472.26	468.58	468.58	33.68	21.82	21.82	282871.05	94549.87	94549.87	1.89	2.88	2.88	68.54	85.42	85.42
6	493.08	498.92	477.08	40.46	39.96	33.03	419566.58	391723.98	248273.95	1.53	1.30	1.37	66.50	51.95	54.07
7	1279.07	751.08	751.08	18.79	21.29	21.29	34975.95	63536.28	63536.28	16.39	9.45	9.45	308.02	201.20	201.20
8	585.16	42.11	454.51	26.88	3.50	12.23	164491.76	3270.42	30114.00	2.56	10.18	4.59	80.74	55.04	56.11
9	641.52	705.02	520.05	24.11	19.07	23.64	93952.40	86285.24	92994.37	3.88	13.67	2.67	127.33	260.59	78.32
10	671.28	759.80	520.05	45.99	36.17	23.64	437671.37	241831.53	92994.37	2.58	3.95	2.67	118.43	152.89	78.32
11	566.03	708.91	708.91	43.18	61.62	61.62	499488.28	961658.82	961658.82	1.27	1.82	1.82	58.67	112.30	112.30
13	644.32	504.75	483.72	67.53	23.68	29.80	1367801.74	113438.68	202385.10	0.91	1.23	1.42	73.60	39.58	43.24
14	648.74	527.58	627.43	25.22	27.84	36.68	114318.52	150379.03	360256.06	2.41	2.07	1.96	76.81	57.66	72.05
15	580.10	542.17	542.17	56.53	43.43	43.43	860413.27	489216.65	489216.65	1.02	1.29	1.29	57.83	57.14	57.14

NJ: normalised jerk (smoothness)
TPA: Time/Path Accuracy

4.4.7.3.5 Tracing task: testing for a Normal distribution

Explorations of the data's distribution suggested that baseline scores were significantly different from a Normal distribution. The skewness *z*-values were all well outside the 1.96 to -1.96 range. The Kologorov-Smirnov Test and the Shapiro-Wilk Test were significant across all baseline outcomes scores except Path Length Time, as shown in Table 4-26, and Normalised Jerk, for which only the Shapiro-Wilk Test showed significance; the Shapiro-Wilk Test is the most sensitive, however (Field, 2009).

Table 4-24. Results of testing for Normal distribution of Tracing results.

		Kolmogorov-Smirnov			**Shapiro-Wilk**		
		Statistic	**df**	***p*-value**	**Statistic**	**df**	***p*-value**
Path Length	Baseline	0.268	14	0.007	0.759	14	0.002
	Six weeks	0.223	14	0.058	0.806	14	0.006
	Twelve weeks	0.195	14	0.153	0.906	14	0.137
Path Length Time	Baseline	0.177	14	0.200	0.929	14	0.292
	Six weeks	0.149	14	0.200	0.96	14	0.722
	Twelve weeks	0.161	14	0.200	0.9	14	0.113
Normalised Jerk	Baseline	0.21	14	0.096	0.772	14	0.002
	Six weeks	0.255	14	0.014	0.754	14	0.001
	Twelve weeks	0.243	14	0.025	0.726	14	0.001
Path Accuracy	Baseline	0.372	14	>0.001	0.525	14	>0.001
	Six weeks	0.286	14	0.003	0.776	14	0.003
	Twelve weeks	0.281	14	0.004	0.791	14	0.004
TPA	Baseline	0.296	14	0.002	0.646	14	>0.001
	Six weeks	0.237	14	0.032	0.807	14	0.006
	Twelve weeks	0.312	14	0.001	0.806	14	0.006

df: degrees of freedom
TPA: Time/Path Accuracy

The data were log transformed (Field, 2009) and histograms generated for each variable. All histograms showed acceptable distribution for parametric statistical operations, and a mixed design ANCOVA was therefore performed including the covariates of gender, age, arm disability (MACS) and use of home computer games.

4.4.7.3.6 Mixed design ANCOVA results for Tracing scores

The results of the repeated-measures (within-participants) ANCOVA for the Tracing parameters are given in given in Table 4-25 below. The results reveal that there were no differences in upper limb kinematics between baseline, six and twelve weeks e.g. for any

Tracing parameters e.g. for Path Length the ANCOVA result was F(2,16) = 0.303, *p*=0.633, adjusting for the covariates of gender, age, arm disability (MACS) and use of home computer games. These results suggest that botulinum treatment made no difference to the children's upper limb kinematics for any of the Tracing parameters across time points. There was no interaction effect between time points and allocation e.g. for Path length F(2,16)=0.998, *p*=0.359. These results show that use of the CAAR games had no influence on the children's upper limb kinematics as measured by the Tracing task.

Table 4-25. Results of within-participants (repeated-measures) ANCOVA for Tracing scores showing significance of comparisons between time points.

	Sum of Squares	df	Mean Square	F	*p*-value
Path Length					
Time points	0.015	2	0.013	0.303	0.633
Time points against allocation	0.050	2	0.042	0.998	0.359
Error (time points)	0.402	16	0.042		
Path Length Time					
Time points	0.003	2	0.002	0.083	0.921
Time points against allocation	0.045	2	0.022	1.066	0.367
Error (time points)	0.336	16	0.021		
Normalised Jerk (smoothness)					
Time points	0.039	2	0.019	0.189	0.830
Time points against allocation	0.209	2	0.104	1.011	0.386
Error (time points)	1.654	16	0.103		
Path Accuracy					
Time points	0.043	2	0.022	1.043	0.375
Time points against allocation	0.001	2	0.000	0.017	0.983
Error (time points)	0.332	16	0.021		
Time x Path Accuracy (TPA)					
Time points	0.049	2	0.025	1.469	0.260
Time points against allocation	0.029	2	0.015	0.864	0.440
Error (time points)	0.269	16	0.017		

df: degrees of freedom
F: F statistic

There were no interaction effects between any covariate across time points.

4.4.7.3.7 Tracing task: between-groups analyses

The result of the between-groups ANCOVA for the Tracing task scores is given in Table 4-26 below, and reveals that there was no difference between groups at each time point for any other Tracing task parameter e.g. Path Length F(1,8)= 1.474 *p*=0.259, suggesting that use of the CAAR gaming device made no impact on the CAAR group's upper limb kinematics.

Table 4-26.Results of between-groups ANCOVA on Tracing task scores.

	Sum of Squares	df	Mean Square	F	*p*-value
Path Length					
Allocation (CAAR or control)	0.042	1	0.042	1.474	0.259
Error	0.229	8	0.029		
Path Length Time					
Allocation (CAAR or control)	0.003	1	0.003	0.76	0.790
Error	0.356	8	0.044		
Normalised Jerk (smoothness)					
Allocation (CAAR or control)	1.471	1	1.471	1.979	0.197
Error	1.906	8	0.238		
Path Accuracy					
Allocation (CAAR or control)	0.060	1	0.060	0.622	0.453
Error	0.776	8	0.097		
Time x Path Accuracy (TPA)					
Allocation (CAAR or control)	0.104	1	2.269	0.170	
Error	0.366	8	0.046		

df: degrees of freedom
F: F statistic

4.4.8 Routine NHS treatment and commercial games use

Of the diaries that were passed to the parents for the purpose of collecting details of any rehabilitation therapy received from their therapists, or additional activity and use of commercial computer games at home, only three were returned. Parents reported that they had either mislaid the diary or had not had time to complete them. The amount of time

spent on these different activities is given in Table 4-29. Additional activities that are likely to be beneficial for reducing activity limitation are also given in Table 4-29 for each of the children.

Table 4-27. Times and details of additional rehabilitation, activities and use of computer games.

	Participants whose activities were recorded in diaries		
	01	**04**	**06**
PT and OT exercises (hours over 6 weeks)			
Daily stretches	10 hours 35 mins		1 hour 55 mins
Home exercise programme		8 hours	
Use of computer games (hours over 6 weeks)	22 hours 55 mins	34 hours 10 mins	10 hours 10 mins
Other beneficial activity (hours over 6 weeks)	16 hours 5 mins	28 hours 5 mins	22 hours 55 mins
Activities undertaken	Swimming Dancing Boxing Basketball Skipping	Swimming Basketball Badminton Hydrotherapy	Netball Dancing Gymnastics/PE Trampoline Bat and ball activity

4.4.9 Feedback on games and participation in the trial

Of the fifteen participants taking part in the study, three returned questionnaires giving feedback (two participants in the CAAR group and one in the control group).

Children in the CAAR group said that they enjoyed the games at first but stated that they quickly became bored. The most popular game was a cooperative game in which the child played both characters that took turns to achieve the aim. Parents reported that participating in the trial was no problem, and reiterated that their children quickly became disinterested in using the CAAR device. The child and parent in the control group agreed with the CAAR participants that taking part in the study was not inconvenient.

Suggestions for themes in future computer games were:

- A memory game (three endorsements);
- A chasing game (two endorsements);

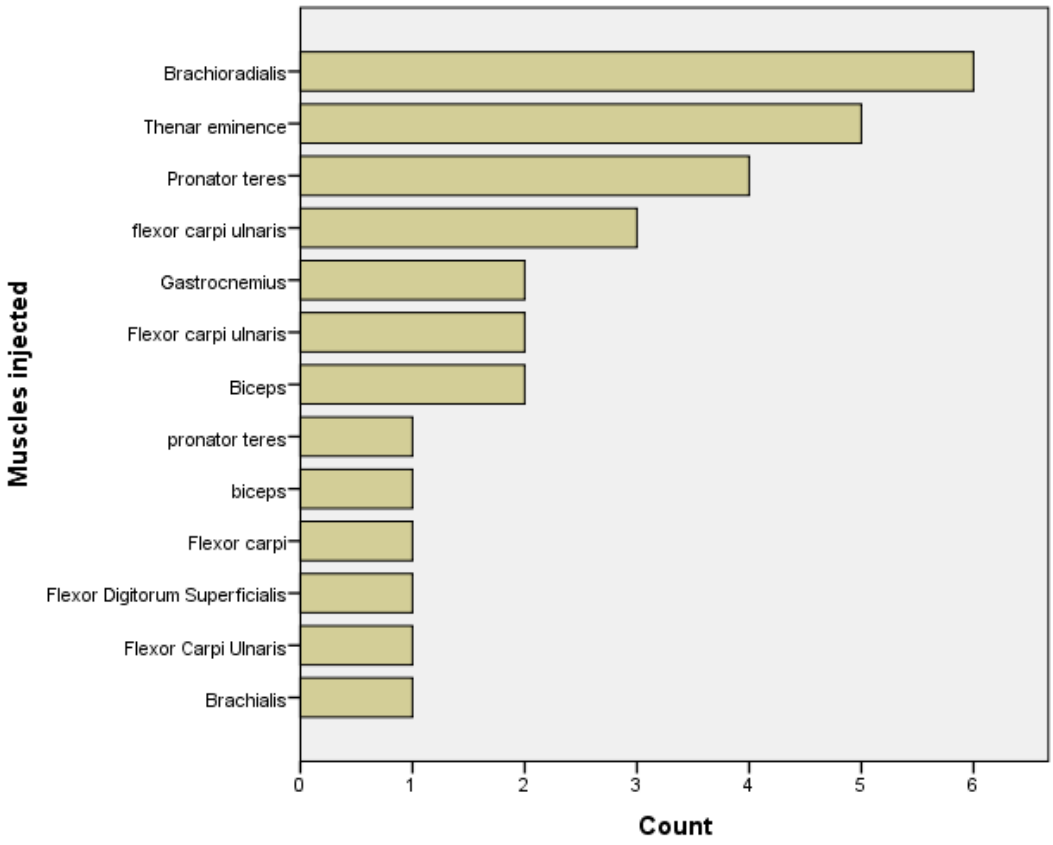
- A puzzle game (two endorsements);
- A racing game (one endorsement);

Spelling games and numbers games received no favourable responses.

4.4.10 Botulinum treatment details

Brachioradialis, the thenar eminence, pronator teres, and flexor carpi ulnaris were the most commonly treated upper limb muscles, as illustrated in Figure 4-21 below. This suggests that spasticity-induced wrist and elbow flexion were the most common muscles causing activity limitation experienced by participants. Thumb-in-palm deformity caused activity limitation in a third of participants.

Figure 4-21. Bar chart illustrating muscles injected across participants.



Details of each participant's botulinum toxin treatment are given in Table 4-30 below.

Table 4-28. Details of participants' treatment with botulinum toxin.

	Gender	Affected arm	Muscles treated with botulinum toxin, and dose in mouse units (MU)							
01	Male	(R) unilateral	Flexor carpi ulnaris	50						
02	Male	(R) unilateral	Thenar eminence	100	Flexor carpi ulnaris	100				
03	Female	(L) unilateral	Thenar eminence	15	Flexor carpi ulnaris	30	Brachioradialis	30		
04	Male	(R) unilateral	Pronator teres	50	Thenar eminence	50	Flexor Carpi Ulnaris	50		
05	Male	(L) unilateral	Flexor carpi ulnaris	50						
06	Female	(R) unilateral	Pronator teres	75						
07	Female	(L) unilateral	Biceps	100	Thenar eminence	50				
08	Female	(R) unilateral	Brachioradialis	100						
09	Male	(R) unilateral	Biceps	35	Brachioradialis	20	Brachialis	15		
10	Male	(L) unilateral	Biceps	30	Brachioradialis	20	Pronator teres	30	Flexor Digitorum Superficialis	20
11	Female	(R) unilateral	Brachioradialis	100						
12	Male	Bilateral	Pronator teres	50						
13	Female	(R) unilateral	flexor carpi ulnaris	50						
14	Male	(L) unilateral	Pronator teres	75						
15	Male	(R) unilateral	Thenar eminence	25	Flexor carpi ulnaris	50				

(L): Left
(R): Right

4.4.11 Adverse events

No adverse events were reported, either for the device or for injuries or harm to participants or their families. A malfunctioning CAAR device wheel was noted by the researchers during the return of a device to the University of Leeds laboratory, and replaced.

Difficulties were experienced during the installation of the device into the home of one child, when final testing of the device caused the main fuse in the residence to operate. The device was replaced the following day, and no further problems were experienced. The device did not malfunction when tested at the University of Leeds laboratory, and no further problems were experienced.

4.5 Discussion

The results of this study suggest that there is no benefit to upper limb activity limitation or arm kinematics from using the CAAR device following botulinum toxin treatment. Furthermore, the results also suggest that botulinum toxin treatment itself did not benefit upper limb activity limitation of the children. However, the low sample size and the low intensity of CAAR use by the children are likely to skew the results substantially so that they are unlikely to represent accurately the true effects either of benefits from using the CAAR device or of botulinum toxin treatment. These limitations might be the reason that the results contradict the results of previous small studies. Firstly, the benefits of botulinum toxin treatment for the upper limb have been established in high quality research (Hoare et al., 2013) and confirmed through a systematic review (Hoare et al., 2010), though these studies highlight the necessity for botulinum toxin to be combined with a structured programme of upper limb rehabilitation. Fasoli et al. (2008) suggest that this structured programme can be replaced by assistive robotic technology. However, their study, and other research on the use of virtual reality or assistive computer technology, is limited to case studies and feasibility studies (Golomb et al., 2010, Gordon et al., 2012, Green and Wilson, 2012, Preston et al., 2014b, Weightman et al., 2011). These studies found that 21 minutes use of technology per day on between 36 and 60 days had the potential to show improvements in upper limb activity limitation (Golomb et al., 2010), but that 12 minutes per day over 13 days showed kinematic improvements but no activity limitation changes (Preston et al., 2014b). There is potential therefore for benefits to upper limb activity limitation from use of this technology as long as there is sufficient intensity of use. The

research highlights both the requirement for investigations into the amount of use of the technology before functional benefits occur and the requirement for standardised and appropriate measures for evaluating outcomes. This Phase II study aimed to build on the results of earlier modelling and feasibility studies that suggested the CAAR device had the potential to improve upper limb activity limitation and upper limb kinematics of children with cerebral palsy (Preston et al., 2014b, Weightman et al., 2011), as described in subsections 2.2.4 and 2.3.2 on pages 55 and 66 respectively.

This thesis tested a single blind randomised controlled design to investigate whether use of the CAAR device enhanced the functional benefits of botulinum toxin treatment of children with cerebral palsy. A number of inclusions to the design of the study, such as blinding of allocation to outcome measure assessors and stratified randomisation using prognostic factors, aimed to prevent bias and overcome flaws of previous trials on CIMT and assistive technology. This thesis also set out to raise the standard of outcome measures available for evaluating activity limitation of children with cerebral palsy. This measure is now provisionally accepted for publication after peer review by a high quality journal and will be available for use in future trials.

The error bars in Figure 4-12 illustrating the effect of botulinum toxin treatment on ABILHAND-kids scores for all participants at six weeks post-injection suggest that there was deterioration in upper limb manual ability. However, the error bars do not take into account the effect of covariates that were predicted *a priori* to influence the outcomes, or the effect of multiple testing. These covariates (manual ability (MACS), age, gender and use of commercial home games) were included in the ANCOVA, which revealed that the differences at each time point were not significant. Nevertheless, only two children on the primary outcome measure and five on the secondary outcome measure showed any improvements, while nine children showed deterioration on the primary outcome measure. Two of the children who deteriorated were the two children who used the games system the most *child 2 and child 13), and one child who improved was the child who barely used the games system at all (child 7). In contrast, on the secondary outcome measure, child 2 showed a clinically significant improvement at six and twelve weeks, while child 7 deteriorated. This illustrates both the possible psychometric deficiencies of these measures and the error associated with drawing conclusions from one or two examples in a manner similar to anecdotal observation. Eighty percent of parents did not return the diaries in which they were asked to document rehabilitation therapy undertaken by their children in

the weeks following rehabilitation, so it is not possible to suggest that these results support the evidence that botulinum toxin is best used as an adjunct to a prospectively arranged and carefully timed rehabilitation programme (Hoare et al., 2013, Hoare et al., 2010).

In contrast to the manual ability scores of the ABILHAND-kids, there was a statistically significant difference in COPM scores between participants at baseline and six weeks, and between baseline and twelve weeks. Although the COPM is recognised as having good responsiveness and although the results suggest that the children improved their performance on the goals selected by their parents, the improvement was not clinically significant, with a change score of less than 2.0 (baseline median score 4.0, six and twelve weeks scores both 4.6). Furthermore, the results of the COPM show that there was no difference between groups at any time point. These results suggest that botulinum toxin treatment has no impact on activity limitation of the upper limb of children with cerebral palsy, and support the finding that there were no benefits from use of the CAAR device. However, the low sample size and minimal use of the device recommend caution with drawing firm conclusions from these results.

Finally the CPKAT kinematic tests revealed that there were no within-group differences between time points in any of the kinematic parameters in either the Pentagram task or the Tracing task. There were no between-groups differences at any time point for the Pentagram but the results did reveal a difference between the CAAR group and control group for the TPA parameter at six weeks. However, there was no within-participants difference between baseline and six weeks, or between groups at baseline, suggesting that this was caused by a non-systematic variation in the scores rather than a systematic variation caused by the CAAR device.

Taken at face value, these results suggest that botulinum toxin does not improve upper limb function of children with cerebral palsy, and that there are no benefits from the additional use of assistive rehabilitation computer gaming devices. The design and diligent conduct of this study was an attempt to address the limitations of previous small studies e.g. the lack of a control group or adequate randomisation; a lack of power calculations; unvalidated, unresponsive or inappropriate outcome measures; non-blinding of assessors; and the use of statistical techniques that were inappropriate for the type of outcome data, distribution of the data or for multiple testing. The results of statistical evaluations suggest that the blinding strategy was successful, and that the stratified randomisation procedure which used a bespoke computer program also achieved its intended aim. Statistical analyses were thorough, and took into account the nature of the data in terms of its type and distribution,

prognostic covariates and multiple testing. The randomisation (minimisation) process balanced the groups satisfactorily with fifteen participants.

Fifteen participants, however, means that the study was underpowered by a large margin, suggesting the strong likelihood of a Type II error, and any results cannot be generalised to the wider population of children with cerebral palsy. The small sample size and the outcome measures used were two limitations of previous studies that this study failed to address, and there are other important factors that are likely to have affected the results to such an extent that no firm conclusions can be drawn about the benefits of botulinum toxin or the CAAR device on upper limb activity limitation of children with cerebral palsy. The two main factors that impacted upon the study are the poor recruitment and the low use by the children of the CAAR device.

4.5.1 Limitations

4.5.1.1 Difficulty in recruiting: small sample size and lack of power

Sample size calculations carried out using the primary outcome measure, the ABILHAND-kids, proposed that 29 children in each group would be an acceptable sample size to achieve 80% power. Although recruitment was initially satisfactory, recruitment slowed when staffing levels at the main recruiting centre (Leeds Teaching Hospitals NHS Trust spasticity clinic) were reduced through unforeseen circumstances. After 21 months, the study had recruited only a quarter of the necessary sample to achieve adequate power. An option was considered to approach other regional spasticity centres but resources restricted researchers' capacity to perform adequate and timely clinic attendance, delivery of devices and follow up visits at children's homes that were some distance away from the research base.

Patient confidentiality prevents inclusion of academic research staff in clinics and other situations where they can approach potential participants. The identification of children who were eligible to take part in the study was necessarily delegated to clerical and clinical staff who, though well briefed on inclusion and exclusion criteria, were unconnected with the study. Offering patients the opportunity to participate in appropriate research is written into the NHS constitution (Department of Health, 2013) and is strongly encouraged by the Department of Health (Department of Health, 2010) and through the NIHR's "It's OK to ask" campaign (National Institute for Health Research, 2013a) but clinicians in rehabilitation teams, though dedicated and professional individuals, have their attention focussed on their caseloads and clinical duties, especially with the increase in those caseloads and other

duties following an implementation of staff redundancies and other austerity measures. This barrier to recruitment impedes good quality research, and is a recognised obstacle (Hewison and Haines, 2006). The NIHR has developed its capability to support studies during the period of this study, and use of its research nurses, who are part of the Trust and have access to clinical areas, might have facilitated the approach to a wider number of children through their attendance at clinics and through meetings with clinic consultants and a review of clinic letters and appointments.

Potentially, the use of clinical trial agreements (CTA) would have helped with ensuring as wide an approach as possible to eligible children, potentially having the effect of giving the clinicians stakeholder status in the study, because they would be named as a Principal Investigator in the CTA. CTAs include an estimate of how many participants that the site expects to recruit. Trust Research and Development (R&D) departments are increasingly monitored on their accrual of participants when participating in NIHR-funded and portfolio-adopted studies. These factors might have increased the incentive and commitment to the recruitment process of clinicians. Against this, however, is the consideration that a Principal Investigator would require Good Clinical Practice (GCP) training, and the submission of a research CV. These additional demands on clinicians were intentionally avoided when designing the methodology, based on the experience of the ChARM study, to avoid the chances of the rehabilitation team managers rejecting their clinicians' participation in the study. In future studies, this might be overcome by funding clinicians' time spent in clinics and attending appropriate research training.

The parents of two children refused to participate because the device was too large to comfortably fit within their homes. This is 10% of all participants approached, and an understandable problem when one considers the small size of back-to-back terraces, stone-built cottages and modern homes. A redesign of the device has been completed, which has had the added advantage of making the device more easily deployed by a single researcher.

The spasticity clinic at the Leeds Teaching Hospitals NHS Trust is one of the earliest to implement botulinum toxin treatment, and has only recently initiated its use in reducing upper limb activity limitation of children with cerebral palsy. The treatment of upper limb spasticity with botulinum toxin is more complex and requires greater clinical experience than the lower limb, and this experience has been developed at pioneering regional centres like the Leeds Teaching Hospitals NHS Trust. The possibility exists therefore that this research is ahead of the clinical experience and capability of the regional hospitals outside of Leeds.

In a future multicentre trial two (or more) further regional centres has the potential to address the slow recruitment and provide a buffer for staff diffuclties at other large centres.

4.5.1.2 Usage of the CAAR games system

The purpose of the CAAR device is to engage children to undertake intensive and repetitive reach-retrieve movements of their impaired upper limb, for which variable assistance is provided for children with more impaired movement. A preliminary review of the CAAR usage data revealed that the average daily use of the CAAR games device by the children in this study was unlikely to have a beneficial effect on arm limitation or arm kinematics. Our school-based feasibility study of the CAAR system (Preston et al., 2014b) found no functional benefits in a similar demographic of children following use of the CAAR system for a median of almost twelve minutes per day, with eight out of the eleven children using the device for more than 20 days out of 40 days. In the home-based feasibility study, eighteen children played for a median of more than an hour a day and showed a statistically-significant difference in COPM scores (although a clinical difference was not quite achieved). In this study, five of the eight children played the CAAR device for less than seven minutes per day, over a mean of eleven days. Only two children played the games system for more than ten minutes, and one of these children played on only three days. These playing times are insufficient to allow any conclusions to be drawn about the possible benefits to upper limb activity limitation of children with cerebral palsy from use of the CAAR device. The lack of engagement with the CAAR device during this home-based study contrasts with the previous home-based and school-based studies. Since this study started Geerdink et al. (2013) have suggested that children older than five years of age require more than 54 hours of unilateral training, at nine hours per week, to reach maximum activity capacity. This is substantially more than that achieved in this study.

There are a number or possible reasons for the disengagement with the games device. The children who participated in the first home-based study expressed a strong desire for a multi-player games system, and this finding was supported by the children who participated in the school-based study. In the current study, children in the CAAR group quickly became bored with the games; this obstacle to engagement and intense practice might have been delayed or prevented by the opportunity to play the games with family and friends. However, this option was unavailable because the MHRA would permit only participants allocated to the CAAR group to use the CAAR device. An online hub or a network similar to that described by Golomb et al. (2010) is being developed that will provide the facility for participants to play each other in real time or as ‘virtual’ opponents, when children play

online and their results are stored and replayed for the opponent when they next log in prior to taking their turn. Secondly, the potential exists that introducing each game in turn after a set period, as was the case in the school-based study described in subsection 2.3.2 on page 66 (Preston et al., 2014b), would have maintained interest in playing the device for a longer period and might have resulted in increased game play and therapeutic movement. This might have increased the game play of the four children who used the games for less than three weeks altogether, with all their use taking place in the first three weeks, and might have increased the game play of the four children who used the games device weekly. There is also the potential that over a six week period these four simple games were insufficient to maintain the children's interest, especially when they were competing against commercial games (80% of the CAAR group children also had commercial games systems), although it is notable that one of the two children without a commercial games system used the CAAR device the least (child 7).

The difficulty in communicating with parents for follow up visits and the low return rate of questionnaires and diaries suggests a lack of support by parents, although the additional pressure of participation in this study on the families is recognised. This pressure was increased by the requirement for use of the CAAR device to be under the supervision of parents, as requested by the MHRA. The requirement for a strong element of encouragement and support from a supervising adult was recognised in our earlier NIHR-funded K005 study (Preston et al., 2014b).

The target of half an hour daily was carefully and intentionally explained to parents as the minimum that would be expected to show benefits to a child's upper limb function, and they were in no doubt as the experimental nature of the study. In hindsight, however, it might have been better to stress that this was an experimental *rehabilitation* device, rather than a games device, that was being evaluated as a supplement to traditional rehabilitation exercises that were essential for the full benefits of botulinum toxin to be realised: in other words, suggesting to parents that thirty minutes of playing on the games device would be a much better option for the children than thirty minutes of stretches and standard rehabilitation exercises, and emphasising that active parental encouragement and engagement would be essential for the potential rehabilitation benefits to be realised.

4.5.1.3 Outcome measures

The limitations of outcome measures have been described in Chapter 3: Critical review of measures of upper limb functional ability for children with cerebral palsy. The primary outcome measure, the ABILHAND-kids, has been developed using modern psychometric

methods (Rasch analysis) and a number of reviews have identified it as one of two measures which meet acceptable psychometric standards. However, the appraisal suggests that it lacks responsiveness to change in manual ability (activity levels), which might have contributed to the apparent lack of effect. The COPM is responsive and there was a significant (but non-clinical) improvement in its outcome scores but it also has its psychometric limitations. The non-linear nature of its ordinal outcome scores and its outcome score generation can amplify any changes in outcome scores, and there are wide confidence intervals around ordinal outcome scores.

These limitations were recognised before the study began, and steps were taken to address them by developing the ChARM. However, this process took two years longer than anticipated, and the ChARM was not available for use by the beginning of the study. The ABILHAND-kids and the COPM were therefore selected because, in spite of the limitations identified, they are the most appropriate measures available for evaluation of upper limb activity limitation in children with cerebral palsy. The ChARM will be available for sue in future studies.

4.6 Conclusion

The results of this Phase II study suggest that it is an appropriate design for evaluating the benefits of assistive technology combined with botulinum toxin. The results of the study revealed that botulinum toxin did not benefit arm activity limitation of this sample of children with cerebral palsy. The study also revealed that use of the CAAR device had no effect on arm activity limitation or kinematics. However, amongst a number of limitations with the study, two in particular – a lack of use of the CAAR device and a small sample size – cause considerable doubt in the findings so that they cannot be accepted with any degree of confidence. Therefore, this study has not been able to prove or disapprove the hypothesis that use of the CAAR device by children with cerebral palsy enhances the functional benefits of upper limb activity limitation. However, this is the aim of the Phase III multicentre study.

It appears that the support of parents and other carers in encouraging children to undertake sufficient daily use of the CAAR device is essential, and that use of the device is likely to be increased if a dual user system is available. Promoting use of the device as an adjunctive component of a daily rehabilitation programme rather than as a games system might also increase its use.

This study achieved adequate blinding of assessors and successfully employed a randomising procedure using a bespoke minimisation computer application. However, the study was very underpowered so that the suggestion of a Type II error cannot be rejected. The use of CTAs and NIHR support should be considered for the purposes of facilitating better support, recruitment and higher quality research.

The lack of benefit on activity limitation seen in the study sample might also be due to a lack of follow up rehabilitation in the weeks immediately following botulinum toxin treatment.

5 Thesis summary, conclusion and further work

The aim of this thesis was to establish whether a computer-assisted arm rehabilitation games device enhanced the established functional benefits of upper limb botulinum toxin treatment of spasticity in children with cerebral palsy. The thesis was developed from studies taken from various disciplines – methodological studies, psychometric studies and studies on experimental rehabilitation and pharmaceutical approaches – and was based upon our own published work that explored the potential of assistive robotic gaming technology to improve upper limb activity limitation of children with cerebral palsy.

5.1 Summary

5.1.1 Background

Cerebral palsy is a common motor disorder that occurs *in utero* or early infancy. It affects the upper limb function in up to 83% of children with spastic cerebral palsy, the type of cerebral palsy which affects the majority (91%) of children with cerebral palsy. This has a major impact on a child's activities and on their participation in social, school and play situations.

In recent years investigators have explored experimental rehabilitation approaches that use high intensity, repetitive unilateral and bilateral training. These have produced encouraging results, including objective changes on neuro-imaging and activity limitation. The use of virtual reality, computerised assistance and robotic technology that operate by facilitating and promoting these approaches is a more recent development. Combining all these approaches with botulinum toxin treatment of spasticity has also met with success, supporting the evidence of a systematic review that suggests that botulinum toxin is an effective intervention for functional benefits only when used with rehabilitation programmes.

However, the evidence produced by the studies exploring these approaches has been adversely affected by two main problems: experimental weakness, and limitations with outcome measures. This thesis set out to address these weaknesses and limitations. In doing so, it aimed to add to the growing body of experimental evidence for assistive rehabilitation technology to promote reduced activity limitation of children with cerebral

palsy, and to support the efforts of other investigators by providing a psychometrically-sound activity limitation measure that is validated for children with cerebral palsy.

5.1.2 The development of a new measure and evidence for its requirement

The scientific testing of a measure is essential for a number of reasons: to validate it for its intended use in its intended population; to determine that it meets the fundamental principles for linear measurement; to determine that it is unidimensional; and to define other psychometric properties e.g. reliability. This thesis established that many of the measures used in paediatric research did not meet adequate standards of modern psychometric testing. Some outcome measures used inappropriate parametric calculations on non-parametric data, and others produced ordinal outcome scores that this thesis argues are not acceptable for research purposes due to wide confidence intervals and misuse of statistical procedures. Others lacked responsiveness, used items poorly-designed for the purpose or population or were simply resource-consuming to obtain and impractical to use. This thesis also looked at the advantages and disadvantages of traditional and modern psychometric methods, and how these advantages and disadvantages contributed to the strengths and weaknesses of each measure. The appraisal of these measures illustrated the requirement for a new measure of activity limitation for children with cerebral palsy, which was conceptualised and developed over a period of three years and realised as the Children's Arm Rehabilitation Measure (ChARM).

The ChARM's items were carefully and diligently constructed to be valid for children with cerebral palsy and relevant to the ICF-CY before its psychometric properties were evaluated using Rasch analysis, a modern psychometric procedure. The initial Rasch analysis identified a number of psychometric limitations with the ChARM, which was then subjected to a substantial revision and secondary Rasch analysis. Finally, a measure of upper limb activity limitation measure valid for children with cerebral palsy was produced that has acceptable, scientifically-validated psychometric properties and which is free for use to academic researchers and health care professionals.

This thesis also provided a platform to carry out the initial psychometric evaluation of a portable kinematic assessment tool for use with children with cerebral palsy.

5.1.3 Evidence-based development of assistive rehabilitation technology and trial design

In parallel with the studies by other investigators, our multidisciplinary team was developing assistive technology with the aim of investigating its potential for benefiting activity

limitation of children with cerebral palsy. The computer-assisted arm rehabilitation (CAAR) technology used in this thesis was the culmination of two feasibility studies, the first of which aimed to develop assistive powered joysticks and games to promote repetitive high-intensity movements for children to use at home. The second study used the findings of the first study, and the results of focus groups involving children with cerebral palsy, to develop a dual-user device for use in schools. There were three aims: to determine which mode (single or dual-use) was most popular; to evaluate the feasibility of deploying rehabilitation devices into schools; and to evaluate any functional and kinematic changes which might result from daily use of the CAAR device. The results of these trials were promising, although we fully recognised their limitations, such as the lack of a control group, that affected the results of other investigators' studies.

Experimental weaknesses include the use of convenience samples, the lack of a control group, non-blinded assessors and unpowered studies. Feasibility and pilot studies do not necessarily require the use of all these elements, but when a trial omits any of them, the trial results contain a degree of uncertainty. This thesis developed a Phase II trial designed to overcome these experimental flaws. Firstly, a sample size calculation was performed to ensure an adequately powered study and prevention of a Type II error. The design included random allocation into intervention and control groups; blinding of the assessors to that allocation, both of which are essential to minimise bias; and blinding to the allocation process. Prognostic factors which might affect the outcomes were considered and these, combined with the relatively small sample size, guided the trial design to use a minimisation technique that both stratified the participants by the prognostic factors and maintained balanced groups throughout the randomisation process.

5.1.4 Outcomes of RCT investigating CAAR combined with botulinum toxin

The results of the RCT reveal that use of the CAAR device for less than eleven minutes daily has no effect on upper limb activity limitation or kinematics when used after botulinum toxin treatment for spasticity of the upper limb. The study also revealed that botulinum toxin treatment does not benefit upper limb activity limitation.

However, the outcomes of the study are limited by the poor rate of recruitment into the trial and a lack of engagement with the CAAR device. The poor recruitment rate resulted in a badly underpowered study, possibly resulting in a Type II error.

The lack of engagement resulted in an amount of CAAR device use that is very unlikely to have produced any functional changes even if the hypothesis is true. This suggests that the

study cannot provide any evidence that the CAAR device does or does not benefit upper limb activity limitation in children with cerebral palsy when it is used in combination with botulinum toxin. This is also likely to affect the results of the study which suggest that botulinum toxin has no effect on upper limb function, although this finding supports the result of a systematic review that suggests botulinum toxin is only effective for activity limitation when used as an adjunct to a rehabilitation programme.

The lack of engagement with the games and the poor recruitment rate were disappointing, but reasons for these can be suggested. Evidence from our previous feasibility studies suggests that a dual-use CAAR device, on which competitive or collaborative games can be played with companions e.g. school friends motivates children to undertake more game play. This option was prevented by regulatory authorities because of potential health and safety concerns. In hindsight, a delay for each game becoming available to the children in a similar method as the school-based feasibility study, e.g. each game becoming available every seven days in turn, might have promoted greater use. These problems could be overcome by developing virtual opponents, or the facility for online gaming. The potential importance of parental and carer support to promote engagement with rehabilitation activity, including use of assistive technology, cannot be dismissed.

5.2 Conclusion

This results of this Phase II single-blind randomised controlled trial suggest that the assistive robotic games device does not enhance the functional benefits of botulinum toxin treatment. However, the limitations of the study give rise to a degree of doubt in the results such that it is not possible to accept or reject the hypothesis. The study design is fit-for-purpose if used in a Phase III trial which will itself be able to overcome the flaws experienced in this smaller trial.

5.3 Future work and collaborations

5.3.1 New measure of participation restriction for children with cerebral palsy and a new school-based measure of fine hand use

The novel method for development of ChARM items was successful, and opens up the potential for development of items for other measures using the same methodology. Sakzewski et al. (2007) suggest that the inclusion of participation measures in clinical trials is necessary for assessing the wider benefits of experimental interventions, and find that there

are no participation measures for children with cerebral palsy, or any that map to a complete range of ICF-CY categories. The NIHR Doctoral Fellowship that funded this research has enabled a number of collaborations and associations, including membership of the Strategic Research Group (SRG) of the British Academy of Childhood Disability (BACD). The SRG is composed of paediatricians, speech and language therapists, psychologists, occupational therapists, physiotherapists, nurses and podiatrists all of whom have an interest in supporting and carrying out research and that work across the UK. This opens up the potential to quickly and efficiently reach a large sample of children's families for the development of items for the new participation measure.

Clinicians and academics associated with this thesis have expressed frustration that no appropriate measures exist for assessing classroom activity limitations in children with motor impairment. Future applications for research funding will allow this to be addressed quickly using the same methodology.

5.3.2 Robotic technology

The thesis has explored the potential of assistive robotic technology to facilitate motor learning by promoting repetition, enhancing feedback and using "active assist" analogous to therapies used by therapists. A further association brought about through this NIHR Fellowship is with the Institute of Psychological Sciences' PACLab at the University of Leeds. The PACLab is a multidisciplinary team that was responsible for the early work on CKAT, and carries out various investigations into how humans interact with their environment. Future research will involve collaborations with PACLab to improve and develop the robotic devices and to explore the feasibility of deploying them to physiotherapy clinics for group work and for providing OT intervention in school settings. The findings of this thesis will provide guidance for these developing technologies.

6 List of References

AARTS, P., JONGERIUS, P., GEERDINK, Y. & GEURTS, A. 2009. Validity and reliability of the VOAA-DDD to assess spontaneous hand use with a video observation tool in children with spastic unilateral cerebral palsy. *BMC Musculoskelet Disord.,* 10**,** 145. doi 10.1186/1471-2474-10-145.

AARTS, P. B., JONGERIUS, P. H., GEERDINK, Y. A., VAN LIMBEEK, J. & GEURTS, A. C. 2010. Effectiveness of Modified Constraint-Induced Movement Therapy in Children With Unilateral Spastic Cerebral Palsy: A Randomized Controlled Trial. *Neurorehabilitation and Neural Repair,* 24**,** 509-518.

AARTS, P. B., JONGERIUS, P. H., GEERDINK, Y. A., VAN LIMBEEK, J. & GEURTS, A. C. 2011. Modified Constraint-Induced Movement Therapy combined with Bimanual Training (mCIMT–BiT) in children with unilateral spastic cerebral palsy: How are improvements in arm-hand use established? *Research in Developmental Disabilities,* 32**,** 271-279.

AARTS, P. B. M., JONGERIUS, P. H., AARTS, M. A. G., VAN HARTINGSVELDT, M. J., ANDERSON, P. G. & BEUMER, A. 2007. A pilot study of the Video Observations Aarts and Aarts (VOAA): a new software program to measure motor behaviour in children with cerebral palsy. *Occupational Therapy International,* 14**,** 113-122.

AISEN, M. L., KERKOVICH, D., MAST, J., MULROY, S., WREN, T. A. L., KAY, R. M. & RETHLEFSEN, S. A. 2011. Cerebral palsy: clinical care and neurological rehabilitation. *The Lancet Neurology,* 10**,** 844-852.

ALBERMAN, E. V. A. & MUTCH, L. 2007. Commentary on the revised versions of the definition and classification of cerebral palsy. *Developmental Medicine & Child Neurology,* 49**,** 32-32.

ALTMAN, D. & BLAND, M. 2005. Treatment allocation by minimisation. *BMJ,* 330**,** 843

ALTMAN, D. G. 1996. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ,* 313**,** 570-571.

ALTMAN, D. G. & BLAND, J. M. 1999a. How to randomise. *BMJ,* 319**,** 703-704.

ALTMAN, D. G. & BLAND, J. M. 1999b. Treatment allocation in controlled trials: why randomise? *BMJ,* 318**,** 1209-1209.

ALTMAN, D. G. & SCHULZ, K. F. 2001. Concealing treatment allocation in randomised trials. *BMJ,* 323**,** 446-447.

ANDERSEN, J. C., MAJNEMER, A., O'GRADY, K. & GORDON, A. M. 2013. Intensive Upper Extremity Training for Children with Hemiplegia: From Science to Practice. *Seminars in Pediatric Neurology,* 20**,** 100-105.

ANTTILA, H., AUTTI-RAMO, I., SUORANTA, J., MAKELA, M. & MALMIVAARA, A. 2008. Effectiveness of physical therapy interventions for children with cerebral palsy: A systematic review. *BMC Pediatrics,* 8**,** 14.

ARNOULD, C., PENTA, M., RENDERS, A., THONNARD, J. & L 2004. ABILHAND-Kids: A measure of manual ability in children with cerebral palsy. *Neurology,* 63**,** 1045-1052.

ARNOULD, C., VANDERVELDE, L., BATCHO, C., PENTA, M. & THONNARD, J. 2012. Can manual ability be measured with a generic ABILHAND scale? A cross-sectional study conducted on six diagnostic groups. *BMJ Open.,* 2(6). pii**,** e001807. doi 10.1136/bmjopen-2012-001807. Print 2012.

AZAULA, M., MSALL, M. E., BUCK, G., TREMONT, M. R., WILCZENSKI, F. & ROGERS, B. T. 2000. Measuring functional status and family support in older school-aged children with cerebral palsy: comparison of three instruments. *Arch Phys Med Rehabil,* 81**,** 307-11.

BARNES, D., LINTON, J., SULLIVAN, E., BAGLEY, A., OEFFINGER, D., ABEL, M., DAMIANO, D., GORTON, G., NICHOLSON, D., ROMNESS, M., ROGERS, S. & TYLKOWSKI, C. 2008. Pediatric outcomes data collection instrument scores in ambulatory children with cerebral palsy: an analysis by age groups and severity level. *Journal of pediatric orthopedics,* 28**,** 97-102.

BARNES, M. 2003. Botulinum toxin--mechanisms of action and clinical use in spasticity. *J Rehabil Med***,** 56-9.

BAX, M. C. 1964. TERMINOLOGY AND CLASSIFICATION OF CEREBRAL PALSY. *Dev Med Child Neurol,* 6**,** 295-7.

BECKUNG, E. & HAGBERG, G. 2002. Neuroimpairments, activity limitations, and participation restrictions in children with cerebral palsy. *Dev Med Child Neurol,* 44**,** 309-16.

BLAIR, E. & WATSON, L. 2006. Epidemiology of cerebral palsy. *Seminars in Fetal and Neonatal Medicine,* 11**,** 117-125.

BLAIR, E. V. E., BADAWI, N. A. & WATSON, L. 2007. Definition and classification of the cerebral palsies: the Australian view. *Developmental Medicine & Child Neurology,* 49**,** 33-34.

BLAND, J. M. & ALTMAN, D. G. 1995. Multiple significance tests: the Bonferroni method. *BMJ,* 310**,** 170.

BOBATH, K. & BOBATH, B. 1984. The neuro-developmental treatment. *In:* SCRUTTON, D. (ed.) *Management of the Motor Disorders of Children with Cerebral Palsy.* 1 ed. Philadephia: JB Lippincott.

BOND, T. G. 2003. Validity and assessment: a Rasch measurement perspective. *Metodología de las Ciencias del Comportamiento,* 5**,** 179 - 194.

BOND, T. G. & FOX, M. C. 2001. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences,* Mawah, New Jersey, Lawrence Erlbaum Associates, Inc.

BONNIER, B., ELIASSON, A. & KRUMLINDE-SUNDHOLM, L. 2006. Effects of constraint-induced movement therapy in adolescents with hemiplegic cerebral palsy: a day camp model. *Scand J Occup Ther.,* 13**,** 13-22.

BORGGRAEFE, I., SCHAEFER, J. S., KLAIBER, M., DABROWSKI, E., AMMANN-REIFFER, C., KNECHT, B., BERWECK, S., HEINEN, F. & MEYER-HEIM, A. 2010. Robotic-assisted treadmill therapy improves walking and standing performance in children and adolescents with cerebral palsy. *European Journal of Paediatric Neurology,* 14**,** 496-502.

BOURKE-TAYLOR, H. 2003. Melbourne Assessment of Unilateral Upper Limb Function: construct validity and correlation with the Pediatric Evaluation of Disability Inventory. *Developmental Medicine and Child Neurology,* 45**,** 92-96.

BOVOLENTA, F. G., M. CLERICI, P. AGOSTI, P. FRANCESCHINI, M. 2009. Robot therapy for functional recovery of the upper limbs: A pilot study on patients after stroke. *Journal of Rehabilitation Medicine,* 41.

BOYD, R., MITCHELL, L., ZIVIANI, J., BILDE, P., KLIIM-DUE, M., RASMUSSEN, B. & NIELSEN, J. 2012. Move it to improve it (Mitii)-Feasibility of a novel web-based therapy for children and adolescents with cerebral palsy. *Developmental Medicine and Child Neurology. Conference: 6th Biennial Conference of the Australasian Academy of Cerebral Palsy and Developmental Medicine Brisbane, QLD Australia. Conference Start,* 54.

BOYD, R., ZIVIANI, J., SAKZEWSKI, L., MILLER, L., BOWDEN, J., CUNNINGTON, R., WARE, R., GUZZETTA, A., AL MACDONELL, R., JACKSON, G., ABBOTT, D. & ROSE, S. 2013. COMBIT: protocol of a randomised comparison trial of COMbined modified constraint induced movement therapy and bimanual intensive training with distributed model of standard upper limb rehabilitation in children with congenital hemiplegia. *BMC Neurology,* 13**,** 68.

BOYD, R. N., MORRIS, M. E. & GRAHAM, H. K. 2001. Management of upper limb dysfunction in children with cerebral palsy: a systematic review. *European Journal of Neurology,* 8**,** 150-166.

BRUININKS, R. H. & BRUININKS, B. D. 2005. Bruininks-Oseretsky Test of Motor Proficiency: Examiners Manual 2ed. Circle Pines, MN: AGS Publishing.

CAMPBELL, M., FITZPATRICK, R., HAINES, A., KINMONTH, A. L., SANDERCOCK, P., SPIEGELHALTER, D. & TYRER, P. 2000. *Framework for design and evaluation of complex interventions to improve health*.

CANS, C., DOLK, H., PLATT, M. J., COLVER, A., PRASAUSK1ENE, A. & RÄGELOH-MANN, I. K. 2007. Recommendations from the SCPE collaborative group for defining and classifying cerebral palsy. *Developmental Medicine & Child Neurology,* 49**,** 35-38.

CARNAHAN, K., ARNER, M. & HAGGLUND, G. 2007. Association between gross motor function (GMFCS) and manual ability (MACS) in children with cerebral palsy. A population-based study of 359 children. *BMC Musculoskeletal Disorders,* 8**,** 50.

CENTRE FOR REVIEWS AND DISSEMINATION 2013. Management of upper limb dysfunction in children with cerebral palsy: a systematic review (Structured abstract). *Database of Abstracts of Reviews of Effects,* 1**,** 1.

CHARLES, J. & GORDON, A. M. 2006. Development of hand-arm bimanual intensive training (HABIT) for improving bimanual coordination in children with hemiplegic cerebral palsy. *Developmental Medicine & Child Neurology,* 48**,** 931-936.

CHIEN, C. & BOND, T. 2009. Measurement properties of fine motor scale of Peabody developmental motor scales-second edition: a Rasch analysis. *Am J Phys Med Rehabil.,* 88**,** 376-86. doi 10.1097/PHM.0b013e318198a7c9.

CHIEN, C. & BROWN, T. 2012. Construct validity of the Children's Hand-Skills ability Questionnaire (CHSQ) in children with disabilities: a Rasch analysis. *Res Dev Disabil.,* 33**,** 1242-53. doi 10.1016/j.ridd.2012.02.023. Epub 2012 Mar 22.

CHIEN, C., BROWN, T. & MCDONALD, R. 2010. Examining content validity and reliability of the Assessment of Children's Hand Skills (ACHS): a preliminary study. *Am J Occup Ther.,* 64**,** 756-67.

CHIEN, C., BROWN, T. & MCDONALD, R. 2011a. Cross-cultural validity of a naturalistic observational assessment of children's hand skills: a study using Rasch analysis. *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine,* 43**,** 631-637.

CHIEN, C., W, BROWN, T. & MCDONALD, R. 2011b. Rasch analysis of the assessment of children's hand skills in children with and without disabilities. *Research in Developmental Disabilities,* 32**,** 253-261.

CHOUDHARY, A., GULATI, S., KABRA, M., SINGH, U. P., SANKHYAN, N., PANDEY, R. M. & KALRA, V. 2013. Efficacy of modified constraint induced movement therapy in improving upper limb function in children with hemiplegic cerebral palsy: A randomized controlled trial. *Brain and Development,* 35**,** 870-876.

COLVER, A. 2007. Classification of cerebral palsy: paediatric perspective. *Developmental Medicine & Child Neurology,* 49**,** 15-16.

CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. & PETTICREW, M. 2008. *Developing and evaluating complex interventions: the new Medical Research Council guidance*.

CULMER, P., LEVESLEY, M., MON-WILLIAMS, M. & WILLIAMS, J. 2009. A new tool for assessing human movement: the Kinematic Assessment Tool. *J Neurosci Methods.,* 184**,** 184-92.

CUSICK, A., LANNIN, N. A. & LOWE, K. 2007. Adapting the Canadian Occupational Performance Measure for use in a paediatric clinical trial. *Disability & Rehabilitation,* 29**,** 761-6.

CUSICK, A., VASQUEZ, M., KNOWLES, L. & WALLEN, M. 2005. Effect of rater training on reliability of Melbourne Assessment of Unilateral Upper Limb Function scores. *Dev Med Child Neurol.,* 47**,** 39-45.

DALTROY, L. H., LIANG, M. H., FOSSEL, A. H. & GOLDBERG, M. J. 1998 The POSNA pediatric musculoskeletal functional health questionnaire: report on reliability, validity, and sensitivity to change. Pediatric Outcomes Instrument Development Group. Pediatric Orthopaedic Society of North America. *J Pediatr Orthop,* 18**,** 561-571.

DAMIANO, D. L. 2006. Activity, activity, activity: rethinking our physical therapy approach to cerebral palsy. *Phys Ther,* 86**,** 1534-40.

DAMIANO, D. L. 2007. Classification of cerebral palsy: clinical therapist's perspective. *Developmental Medicine & Child Neurology,* 49**,** 16-17.

DAMMANN, O. & KUBAN, K. C. 2007. 'Cerebral palsy'— rejected, refined, recovered. *Developmental Medicine & Child Neurology,* 49**,** 17-18.

DAMMANN, O. & LEVITON, A. 1998. Infection remote from the brain, neonatal white matter damage, and cerebral palsy in the preterm infant. *Seminars in Pediatric Neurology,* 5**,** 190-201.

DAVIDS, J. R., PEACE, L. C., WAGNER, L. V., GIDEWALL, M. A., BLACKHURST, D. W. & ROBERSON, W. M. 2006. Validation of the Shriners Hospital for Children Upper Extremity Evaluation (SHUEE) for

Children with Hemiplegic Cerebral Palsy. *The Journal of Bone & Joint Surgery,* 88**,** 326-333.

DAY, S. J. & ALTMAN, D. G. 2000. Blinding in clinical trials and other studies. *BMJ,* 321**,** 504.

DEBUSE, D. & BRACE, H. 2011. Outcome measures of activity for children with cerebral palsy: a systematic review. [Review]. *Pediatric Physical Therapy,* 23**,** 221-31.

DEITZ, J., KARTIN, D. & KOPP, K. 2007. Review of the Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2). *Phys Occup Ther Pediatr,* 27**,** 87-102.

DELUCA, S., ECHOLS, K., LAW, C. & RAMEY, S. 2006. Intensive pediatric constraint-induced therapy for children with cerebral palsy: randomized, controlled, crossover trial. . *J Child Neurol,* 21**,** 931-938.

DEMATTEO, C., LAW, M., RUSSELL, D., POLLOCK, N., ROSENBAUM, P. & WALTER, S. 1993. The reliability and validity of the Quality of Upper Extremity Skills Test. *Physical and Occupational Therapy in Pediatrics,* 13**,** 1-18.

DEPARTMENT OF EDUCATION 2003. Every Child Matters: Change For Children. Department of Education.

DEPARTMENT OF EDUCATION 2004. Removing Barriers To Achievement: The Government's Strategy for SEN. Department of Education.

DEPARTMENT OF HEALTH 2010. Embedding Health Research: National Institute for Health Research Annual Report 2009/10. National Institute for Health Research.

DEPARTMENT OF HEALTH 2013. The NHS Constitution. Department of Health.

DEUTSCH, J. E., BORBELY, M., FILLER, J., HUHN, K. & GUARRERA-BOWLBY, P. 2008. Use of a Low-Cost, Commercially Available Gaming Console (Wii) for Rehabilitation of an Adolescent With Cerebral Palsy. *Physical Therapy,* 88**,** 1196-1207.

DEVELLIS, R. F. 2006. Classical test theory. *Med Care,* 44**,** S50-9.

DRUIN, A. 2002. The role of children in the design of new technology. *Behaviour and Information Technology,* 21**,** 1-25.

ELBERT, T., PANTEV, C., WIENBRUCH, C., ROCKSTROH, B. & TAUB, E. 1995. Increased cortical representation of the fingers of the left hand in string players. *Science,* 270**,** 305-7.

ELIASSON, A. C. 2005. Improving the use of hands in daily activities: aspects of the treatment of children with cerebral palsy. *Phys Occup Ther Pediatr,* 25**,** 37-60.

ELIASSON, A. C., KRUMLINDE-SUNDHOLM, L., ROSBLAD, B., BECKUNG, E., ARNER, M., OHRVALL, A. M. & ROSENBAUM, P. 2006. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol,* 48**,** 549-54.

ELIASSON, A. C., KRUMLINDE-SUNDHOLM, L., SHAW, K. & WANG, C. 2005. Effects of constraint-induced movement therapy in young children with hemiplegic cerebral palsy: an adapted model. *Dev Med Child Neurol,* 47**,** 266-75.

ELLIS, D. 1982. Joey Deacon: a suitable case for labelling? *Dev Med Child Neurol,* 24**,** 485-8.

FASOLI, S., FRAGALA-PINKHAM, M., HUGHES, R., KREBS, H., HOGAN, N. & STEIN, J. 2008. Robotic Therapy and Botulinum Toxin Type A: A Novel Intervention Approach for Cerebral Palsy. *American Journal of Physical Medicine and Rehabilitation,* 87**,** 1022-5.

FASOLI, S. E., FRAGALA-PINKHAM, M., HUGHES, R., HOGAN, N., STEIN, J. & KREBS, H. I. 2010. Upper limb robot-assisted therapy: A new option for children with hemiplegia. *Technology and Disability,* 22**,** 193-198.

FIELD, A. 2009. *Discovering statistics using SPSS*, Sage publications.

FISHER, A. G. & MERRITT, B. K. 2012. Conceptualizing and developing the AMPS within a framework of modern objective measurement. *In:* FISHER, A. G. & JONES, K. B. (eds.) *Assessment of Motor and Process Skills. Vol. 1: Development, Standardization, and Administration Manual (revised 7th ed.).* 7 ed. Fort Collins, CO: Three Star Press

FITOUSSI, F., DIOP, A., MAUREL, N., LAASEL EL, M., ILHARREBORDE, B. & PENNECOT, G. F. 2011. Upper limb motion analysis in children with hemiplegic cerebral palsy: proximal kinematic changes after distal botulinum toxin or surgical treatments. *J Child Orthop,* 5**,** 363-70.

FITOUSSI, F., DIOP, A., MAUREL, N., LAASSEL, E. M. & PENNECOT, G. F. 2006. Kinematic analysis of the upper limb: a useful tool in children with cerebral palsy. *J Pediatr Orthop B,* 15**,** 247-56.

FITZGERALD, M. J. T., GRUENER, G. & MTUI, E. 2012. *Clinical Neuroanatomy and Neuroscience,* China, Saunders Elsevier.

FLUET, G. G., QIU, Q., KELLY, D., PARIKH, H. D., RAMIREZ, D., SALEH, S. & ADAMOVICH, S. V. 2010. Interfacing a haptic robotic system with complex virtual environments to treat impaired upper extremity motor function in children with cerebral palsy. *Dev Neurorehabil,* 13**,** 335-45.

FRASCARELLI, F., MASIA, L., DI ROSA, G., PETRARCA, M., CAPPA, P. & CASTELLI, E. 2009. Robot-mediated and clinical scales evaluation after upper limb botulinum toxin type A injection in children with hemiplegia. *J Rehabil Med,* 41**,** 988-94.

GALEA, M. 2004. Neural plasticity and learning: the potential for change. *In:* SCRUTTON, D., DAMIANO, D. L. & MAYSTON, M. (eds.) *Management of the motor disorders of children with cerebral palsy.* Cambridge: Cambridge University Press.

GARROW, J. S. & WEBSTER, J. 1985. Quetelet's index (W/H2) as a measure of fatness. *Int J Obes,* 9**,** 147-53.

GEERDINK, Y., AARTS, P. & GEURTS, A. C. 2013. Motor learning curve and long-term effectiveness of modified constraint-induced movement therapy in children with unilateral cerebral palsy: A randomized controlled trial. *Research in Developmental Disabilities,* 34**,** 923-931.

GEERDINK, Y., LINDEBOOM, R., DE WOLF, S., STEENBERGEN, B., GEURTS, A. C. & AARTS, P. 2014. Assessment of upper limb capacity in children with unilateral cerebral palsy: construct validity of a Rasch-reduced Modified House Classification. *Dev Med Child Neurol,* 56**,** 580-6.

GILMORE, R., SAKZEWSKI, L. & BOYD, R. 2010. Upper limb activity measures for 5- to 16-year-old children with congenital hemiplegia: A systematic review. *Developmental Medicine and Child Neurology,* 52**,** 14-21.

GOLDENBERG, R. L., CULHANE, J. F., IAMS, J. D. & ROMERO, R. 2008. Epidemiology and causes of preterm birth. *Lancet,* 371**,** 75-84.

GOLOMB, M. R., MCDONALD, B. C., WARDEN, S. J., YONKMAN, J., SAYKIN, A. J., SHIRLEY, B., HUBER, M., RABIN, B., ABDELBAKY, M., NWOSU, M. E., BARKAT-MASIH, M. & BURDEA, G. C. 2010. In-home virtual reality videogame telerehabilitation in adolescents with hemiplegic cerebral palsy. *Arch Phys Med Rehabil,* 91**,** 1-8 e1.

GORDON, A. 2007. Measuring 'activity limitation' in individuals with unilateral upper extremity impairments. *Developmental Medicine and Child Neurology,* 49.

GORDON, A. M., CHARLES, J. & WOLF, S. L. 2006. Efficacy of constraint-induced movement therapy on involved upper-extremity use in children with hemiplegic cerebral palsy is not age-dependent. *Pediatrics,* 117**,** e363-73.

GORDON, A. M., SCHNEIDER, J. A., CHINNAN, A. & CHARLES, J. R. 2007. Efficacy of a hand-arm bimanual intensive therapy (HABIT) in children with hemiplegic cerebral palsy: a randomized control trial. *Dev Med Child Neurol,* 49**,** 830-8.

GORDON, C., ROOPCHAND-MARTIN, S. & GREGG, A. 2012. Potential of the Nintendo Wii™ as a rehabilitation tool for children with cerebral palsy in a developing country: a pilot study. *Physiotherapy,* 98**,** 238-242.

GRAHAM, H. K. 2007. Classification of cerebral palsy: the surgeon's perspective. *Developmental Medicine & Child Neurology,* 49**,** 21-23.

GRAHAM, H. K., AOKI, K. R., AUTTI-RAMO, I., BOYD, R. N., DELGADO, M. R., GAEBLER-SPIRA, D. J., GORMLEY, M. E., GUYER, B. M., HEINEN, F., HOLTON, A. F., MATTHEWS, D., MOLENAERS, G., MOTTA, F., GARCIA RUIZ, P. J. & WISSEL, J. 2000. Recommendations for the use of botulinum toxin type A in the management of cerebral palsy. *Gait Posture,* 11**,** 67-79.

GRANT, M. J. & BOOTH, A. 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J,* 26**,** 91-108.

GREAVES, S., IMMS, C., DODD, K. & KRUMLINDE-SUNDHOLM, L. 2010. Assessing bimanual performance in young children with hemiplegic cerebral palsy: a systematic review. *Developmental Medicine & Child Neurology,* 52**,** 413-421.

GREEN, D., SCHERTZ, M., GORDON, A. M., MOORE, A., SCHEJTER MARGALIT, T., FARQUHARSON, Y., BEN BASHAT, D., WEINSTEIN, M., LIN, J. P. & FATTAL-VALEVSKI, A. 2013. A multi-site study of functional outcomes following a themed approach to hand-arm bimanual intensive therapy for children with hemiplegia. *Dev Med Child Neurol,* 55**,** 527-33.

GREEN, D. & WILSON, P. H. 2012. Use of virtual reality in rehabilitation of movement in children with hemiplegia--a multiple case study evaluation. *Disabil Rehabil,* 34**,** 593-604.

GRETHER, J. K., NELSON, K. B., EMERY III, E. S. & CUMMINS, S. K. 1996. Prenatal and perinatal factors and cerebral palsy in very low birth weight infants. *The Journal of Pediatrics,* 128**,** 407-414.

GRIMBY, G., TENNANT, A. & TESIO, L. 2012. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med.,* 44**,** 97-8. doi 10.2340/16501977-0938.

HARVEY, A., ROBIN, J., MORRIS, M. E., GRAHAM, H. K. & BAKER, R. 2008. A systematic review of measures of activity limitation for children with cerebral palsy. *Developmental Medicine and Child Neurology,* 50**,** 190-198.

HEMMINGSSON, H. & JONSSON, H. 2005. An occupational perspective on the concept of participation in the International Classification of Functioning, Disability and Health--some critical remarks. *Am J Occup Ther,* 59**,** 569-76.

HENDERSON, S. & SUGDEN, D. 1992. *Movement Assessment Battery for Children manual,* Sidcup, UK, The Psychological Corporation.

HEWISON, J. & HAINES, A. 2006. Overcoming barriers to recruitment in health research. *BMJ,* 333**,** 300-2.

HIMMELMANN, K., BECKUNG, E., HAGBERG, G. & UVEBRANT, P. 2006. Gross and fine motor function and accompanying impairments in cerebral palsy. *Dev Med Child Neurol,* 48**,** 417-23.

HOARE, B. & IMMS, C. 2004. Upper-limb injections of botulinum toxin-A in children with cerebral palsy: a critical review of the literature and clinical implications for occupational therapists. *Am J Occup Ther.,* 58**,** 389-97.

HOARE, B., IMMS, C., CAREY, L. & WASIAK, J. 2007a. Constraint-induced movement therapy in the treatment of the upper limb in children with hemiplegic cerebral palsy: a Cochrane systematic review. *Clin Rehabil,* 21**,** 675-85.

HOARE, B., IMMS, C., RANDALL, M. & CAREY, L. 2011. Linking cerebral palsy upper limb measures to the International Classification of Functioning, Disability and Health. *J Rehabil Med.,* 43**,** 987-96. doi: 10.2340/16501977-0886.

HOARE, B., IMMS, C., VILLANUEVA, E., RAWICKI, H. B., MATYAS, T. & CAREY, L. 2013. Intensive therapy following upper limb botulinum toxin A injection in young children with unilateral cerebral palsy: a randomized trial. *Dev Med Child Neurol,* 55**,** 238-47.

HOARE, B. J., WALLEN, M. A., IMMS, C., VILLANUEVA, E., RAWICKI, H. B. & CAREY, L. 2010. Botulinum toxin A as an adjunct to treatment in the management of the upper limb in children with spastic cerebral palsy (UPDATE). *Cochrane Database Syst Rev***,** CD003469.

HOARE, B. J., WASIAK, J., IMMS, C. & CAREY, L. 2007b. Constraint-induced movement therapy in the treatment of the upper limb in children with hemiplegic cerebral palsy. *Cochrane Database Syst Rev***,** CD004149.

HOBART, J. C., CANO, S. J., ZAJICEK, J. P. & THOMPSON, A. J. 2007. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *The Lancet Neurology,* 6**,** 1094 -105.

HOBART, J. C., LAMPING, D. L. & THOMPSON, A. J. 1996. Evaluating neurological outcome measures: the bare essentials. *Journal of Neurology, Neurosurgery, and Psychiatry,* 60**,** 127-130.

HOLMEFUR, M., AARTS, P., HOARE, B. & KRUMLINDE-SUNDHOLM, L. 2009. Retest and alternate forms reliability of the Assisting Hand

Assessment. *Developmental Medicine and Child Neurology. Conference: 3rd International Cerebral Palsy Conference Sydney, NSW Australia. Conference Start,* 51.

HOLT, R., WEIGHTMAN, A., GALLAGHER, J., PRESTON, N., LEVESLEY, M., MON-WILLIAMS, M. & BHAKTA, B. 2013. A System in the Wild: Deploying a Two Player Arm Rehabilitation System for Children with Cerebral Palsy in a School Environment. *Journal of Usability Studies,* 8**,** 111-126.

HORTON, M., MARAIS, I. & CHRISTENSEN, K. B. 2013. Dimensionality. *In:* CHRISTENSEN, K. B., KREINER, S. & MESBAH, M. (eds.) *Rasch Models in Health.* London: ISTE Ltd.

HOUSE, J., GWATHMEY, F. & FIDLER, M. 1981. A dynamic approach to the thumb-in palm deformity in cerebral palsy. *J Bone Joint Surg Am.,* 63**,** 216-25.

HOUWINK, A., GEERDINK, Y., STEENBERGEN, B., GEURTS, A. & AARTS, P. 2013. Assessment of upper-limb capacity, performance, and developmental disregard in children with cerebral palsy: Validity and reliability of the revised Video-Observation Aarts and Aarts module: Determine Developmental Disregard (VOAA-DDD-R). *Developmental Medicine and Child Neurology,* 55**,** 76-82.

HÖYSNIEMI, J., HÄMÄLÄINEN, P. & TURKKI, L. 2003. Using peer tutoring in evaluating the usability of a physically interactive computer game with children. *Interacting with Computers,* 15**,** 203-225.

HUANG, H.-H., FETTERS, L., HALE, J. & MCBRIDE, A. 2009. Bound for Success: A Systematic Review of Constraint-Induced Movement Therapy in Children With Cerebral Palsy Supports Improved Arm and Hand Use. *Physical therapy,* 89**,** 1126-1141.

HUTTON, J. L. & PHAROAH, P. O. 2006. Life expectancy in severe cerebral palsy. *Arch Dis Child,* 91**,** 254-8.

IYER, L. V., HALEY, S. M., WATKINS, M. P. & DUMAS, H. M. 2003. Establishing minimal clinically important differences for scores on the pediatric evaluation of disability inventory for inpatient rehabilitation. *Phys Ther,* 83**,** 888-98.

JACOBSSON, B. & HAGBERG, G. 2004. Antenatal risk factors for cerebral palsy. *Best Practice & Research Clinical Obstetrics & Gynaecology,* 18**,** 425-436.

JASPERS, E., DESLOOVERE, K., BRUYNINCKX, H., KLINGELS, K., MOLENAERS, G., AERTBELIEN, E., GESTEL, L. V. & FEYS, H. 2011. Three-dimensional upper limb movement characteristics in children with hemiplegic cerebral palsy and typically developing children. *Research in Developmental Disabilities.,* 32**,** 2283-2294.

JOHNSON, A. 2002. Prevalence and characteristics of children with cerebral palsy in Europe. *Developmental Medicine and Child Neurology,* 44**,** 633 - 640.

JOHNSON, L., RANDALL, M., REDDIHOUGH, D., OKE, L., BYRT, T. & BACH, T. 1994. Development of a clinical assessment of quality of movement for unilateral upper-limb functon. *Developmental Medicine and Child Neurology,* 36**,** 965-973.

KING, N. A., HILLS, A. P. & BLUNDELL, J. E. 2005. High Body Mass Index is not a barrier to physical activity: Analysis of international rugby players' anthropometric data. *European Journal of Sport Science,* 5**,** 73-75.

KIRTON, A. 2013. Modeling developmental plasticity after perinatal stroke: defining central therapeutic targets in cerebral palsy. *Pediatr Neurol,* 48**,** 81-94.

KLINGELS, K., DE, C. P., DESLOOVERE, K., HUENAERTS, C., MOLENAERS, G., VAN, N. I., HUYSMANS, A. & FEYS, H. 2008. Comparison of the Melbourne assessment of unilateral upper limb function and the quality of upper extremity skills test in hemiplegic CP. *Developmental Medicine and Child Neurology,* 50**,** 904-909.

KLINGELS, K., JASPERS, E., VAN, D. W. A., DE, C. P., MOLENAERS, G. & FEYS, H. 2010. A systematic review of arm activity measures for children with hemiplegic cerebral palsy. *Clinical rehabilitation,* 24**,** 887-900.

KNOX, V. & EVANS, A. L. 2002. Evaluation of the functional effects of a course of Bobath therapy in children with cerebral palsy: a preliminary study. *Dev Med Child Neurol,* 44**,** 447-60.

KOMAN, L., SMITH, B., WILLIAMS, R., RICHARDSON, R., NAUGHTON, M., GRIFFIN, L. & EVANS, P. 2013. Upper extremity spasticity in children with cerebral palsy: A randomized, double-blind, placebo-controlled study of the short-term outcomes of treatment with botulinum a toxin. *Journal of Hand Surgery,* 38**,** 435-446.

KOMAN, L., WILLIAMS, R., EVANS, P., RICHARDSON, R., NAUGHTON, M., PASSMORE, L. & SMITH, B. 2008. Quantification of upper extremity function and range of motion in children with cerebral palsy. *Developmental Medicine and Child Neurology,* 50**,** 910-917.

KOMAN, L. A., SMITH, B. P. & SHILT, J. S. 2004. Cerebral palsy. *The Lancet,* 363**,** 1619-1631.

KRAGELOH-MANN, I. & HORBER, V. 2007. The role of magnetic resonance imaging in elucidating the pathogenesis of cerebral palsy: a systematic review. *Dev Med Child Neurol,* 49**,** 144-51.

KREBS, H. I., LADENHEIM, B., HIPPOLYTE, C., MONTERROSO, L. & MAST, J. 2009. Robot-assisted task-specific training in cerebral palsy. *Dev Med Child Neurol,* 51 Suppl 4**,** 140-5.

KRUMLINDE-SUNDHOLM, L., ELIASSON, A. & C 2003. Development of the assisting hand assessment: A Rasch-built measure intended for children with unilateral upper limb impairments. *Scandinavian Journal of Occupational Therapy,* 10**,** 16-26.

KRUMLINDE-SUNDHOLM, L., HOLMEFUR, M., KOTTORP, A. & ELIASSON, A. 2007. The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. *Developmental Medicine and Child Neurology,* 49**,** 259-64.

LAW, M., BAPTISTE, S., CARSWELL, A., MCCOLL, M. A., POLATAJKO, H. & POLLOCK, N. 2005. Canadian Occupational Performance Measure. *In:* THERAPISTS, C. A. O. O. (ed.) 5 ed. Ottawa, ON: CAOT Publications ACE.

LEMMENS, R., TIMMERMANS, A., JANSSEN-POTTEN, Y., SMEETS, R. & SEELEN, H. 2012. Valid and reliable instruments for arm-hand assessment at ICF activity level in persons with hemiplegia: a systematic review. *BMC Neurology,* 12.

LERMAN, J., SULLIVAN, E., BARNES, D. & HAYNES, R. 2005. The Pediatric Outcomes Data Collection Instrument (PODCI) and functional

assessment of patients with unilateral upper extremity deficiencies. *Journal of Pediatric Orthopaedics,* 25**,** 405-407.

LEVITT, S. 2010. *Treatment of Cerebral Palsy and Motor Delay,* Oxford, Wiley-Blackwell.

LIN, K., CHEN, H., CHEN, C., WANG, T., WU, C., HSIEH, Y. & WU, L. 2012. Validity, responsiveness, minimal detectable change, and minimal clinically important change of the Pediatric Motor Activity Log in children with cerebral palsy. *Res Dev Disabil.,* 33**,** 570-7. doi 10.1016/j.ridd.2011.10.003. Epub 2011 Nov 24.

LINACRE, J. 1994. Sample Size and Item Calibration Stability. *Rasch Measurement Transactions,* 7.

LINACRE, J. M. 2000. FIM levels as ordinal categories. *J Outcome Meas,* 4**,** 616-33.

LINDEBOOM, R., SPRANGERS, M. A. & ZWINDERMAN, A. H. 2005. Responsiveness: a reinvention of the wheel? *Health Qual Life Outcomes,* 3**,** 8.

LONGO, L. D. & ASHWAL, S. 1993. William Osler, Sigmund Freud and the evolution of ideas concerning cerebral palsy. *J Hist Neurosci,* 2**,** 255-82.

MACCOBB, S., GREENE, S., NUGENT, K. & O'MAHONY, P. 2005. Measurement and prediction of motor proficiency in children using Bayley infant scales and the Bruininks-Oseretsky test. *Phys Occup Ther Pediatr.,* 25**,** 59-79.

MAJNEMER, A. & MAZER, B. 2004. New directions in the outcome evaluation of children with cerebral palsy. *Seminars in Pediatric Neurology,* 11**,** 11-17.

MASTERS, G. 1982. A rasch model for partial credit scoring. *Psychometrika,* 47**,** 149-174.

MAYSTON, M. J. 2004. Physiotherapy management in cerebral palsy: An update on treatment approaches. *In:* SCRUTTON, D., DAMIANO, D. L. & MAYSTON, M. (eds.) *Management of the motor disorders of children with cerebral palsy.* Cambridge: Cambridge University Press.

MCCARTHY, M. L., SILBERSTEIN, C. E., ATKINS, E. A., HARRYMAN, S. E., SPONSELLER, P. D. & HADLEY-MILLER, N. A. 2002. Comparing reliability and validity of pediatric instruments for measuring health and well-being of children with spastic cerebral palsy. *Developmental Medicine & Child Neurology,* 44**,** 468-476.

MCCULLOUGH, N. & PARKES, J. 2008. Use of the child health questionnaire in children with cerebral palsy: a systematic review and evaluation of the psychometric properties. *J Pediatr Psychol.,* 33**,** 80-90. Epub 2007 Aug 28.

MCDONALD, E. T. & CHANCE, B. 1964. *Cerebral palsy,* Englewood Cliffs, New Jersey, Prentice-Hall.

MCLAUGHLIN, J. F. 2007. Definition of cerebral palsy: clinical perspective. *Developmental Medicine & Child Neurology,* 49**,** 27-28.

MERBITZ, C., MORRIS, J. & GRIP, J. C. 1989. Ordinal scales and foundations of misinference. *Arch Phys Med Rehabil,* 70**,** 308-12.

MEYER-HEIM, A. & VAN HEDEL, H. J. 2013. Robot-assisted and computer-enhanced therapies for children with cerebral palsy: current state and clinical implementation. *Semin Pediatr Neurol,* 20**,** 139-45.

MILLER, F. (ed.) 2007. *Physical therapy of cerebral palsy*: Springer Verlag.

MILLER, G. & CLARK, G. D. 1998. *The Cerebral Palsies: Causes, consequences, and management,* Boston, Butterworth–Heinemann.

MITCHELL, L., ZIVIANI, J., OFTEDAL, S. & BOYD, R. 2012. The effect of virtual reality interventions on physical activity in children and adolescents with early brain injuries including cerebral palsy. *Developmental Medicine & Child Neurology,* 54**,** 667-671.

MITTRACH, R., GRILL, E., WALCHNER-BONJEAN, M., SCHEURINGER, M., BOLDT, C., HUBER, E. O. & STUCKI, G. 2008. Goals of physiotherapy interventions can be described using the International Classification of Functioning, Disability and Health. *Physiotherapy,* 94**,** 150-157.

MORRIS, C. 2007. Definition and classification of cerebral palsy: a historical perspective. *Dev Med Child Neurol Suppl,* 109**,** 3-7.

MORRIS, C. & BARTLETT, D. 2004. Gross Motor Function Classification System: impact and utility. *Dev Med Child Neurol,* 46**,** 60-5.

MUTCH, L., ALBERMAN, E., HAGBERG, B., KODAMA, K. & PERAT, M. V. 1992. Cerebral palsy epidemiology: where are we now and where are we going? *Dev Med Child Neurol,* 34**,** 547-51.

NATIONAL INSTITUTE FOR HEALTH RESEARCH. 2013a. *It's OK to ask - the NIHR's new patient empowerment campaign* [Online]. Available: http://www.ct-toolkit.ac.uk/news/its-ok-to-ask-the-nihrs-new-patient-empowerment-campaign.

NATIONAL INSTITUTE FOR HEALTH RESEARCH. 2013b. *UK Clinical Research Network Study Portfolio* [Online]. Available: http://public.ukcrn.org.uk/Search/StudyDetail.aspx?StudyID=6306 [Accessed 26/03/2013.

NELSON, K. B. & ELLENBERG, J. H. 1986. Antecedents of Cerebral Palsy. *New England Journal of Medicine,* 315**,** 81-86.

NELSON, K. B. & GRETHER, J. K. 1999. Causes of cerebral palsy. *Curr Opin Pediatr,* 11**,** 487-91.

NICHOLS, D. S. & CASE-SMITH, J. 1996. Reliability and Validity of the Pediatric Evaluation of Disability Inventory. *Pediatric Physical Therapy Spring,* 8**,** 15-24.

NOVAK, I., CUSICK, A. & LANNIN, N. 2009. Occupational Therapy Home Programs for Cerebral Palsy: Double-Blind, Randomized, Controlled Trial. *Pediatrics,* 124**,** e606-e614.

NOVAK, I., HINES, M., GOLDSMITH, S. & BARCLAY, R. 2012. Clinical prognostic messages from a systematic review on cerebral palsy. *Pediatrics,* 130**,** e1285-312.

NOVAK, I., MCINTYRE, S., MORGAN, C., CAMPBELL, L., DARK, L., MORTON, N., STUMBLES, E., WILSON, S.-A. & GOLDSMITH, S. 2013. A systematic review of interventions for children with cerebral palsy: state of the evidence. *Developmental Medicine & Child Neurology,* 55**,** 885-910.

NUDO, R. J. 2003. Adaptive plasticity in motor cortex: implications for rehabilitation after brain injury. *J Rehabil Med***,** 7-10.

O'SHEA, M. 2008. Cerebral Palsy. *Seminars in Perinatology,* 32**,** 35-41.

ODDING, E., ROEBROECK, M. & STAM, H. J. 2006. The epidemiology of cerebral palsy: Incidence, impairments and risk factors. *Disability & Rehabilitation,* 28**,** 183 - 191.

OLIVER, M., SAPEY, B. & THOMAS, P. 2012. *Social Work with Disabled People* Palgrave Macmillan.

OTTENBACHER, K., MSALL, M., LYON, N., DUFFY, L., GRANGER, C. & BRAUN, S. 1997. Interrater agreement and stability of the Functional Independence Measure for Children (WeeFIM): use in children with developmental disabilities. *Arch Phys Med Rehabil.,* 78**,** 1309-15.

OTTENBACHER, K., TAYLOR, E., MSALL, M., BRAUN, S., LANE, S., GRANGER, C., LYONS, N. & DUFFY, L. 1996. The stability and equivalence reliability of the functional independence measure for children (WeeFIM). *Dev Med Child Neurol.,* 38**,** 907-16.

OTTENBACHER, K. J., MSALL, M. E., LYON, N., DUFFY, L. C., ZIVIANI, J., GRANGER, C. V., BRAUN, S. & FEIDLER, R. C. 2000. The WeeFIM instrument: its utility in detecting change in children with developmental disabilities. *Arch Phys Med Rehabil,* 81**,** 1317-26.

PALISANO, R., ROSENBAUM, P., WALTER, S., RUSSELL, D., WOOD, E. & GALUPPI, B. 1997. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol,* 39**,** 214-23.

PALISANO, R. J., KOLOBE, T. H., HALEY, S. M., PAX LOWES, L. & JONES, S. L. 1995. Validity of the Peabody Developmental Gross Motor Scale as an Evaluative Measure of Infants Receiving Physical Therapy. *Physical therapy,* 75**,** 939-948.

PALISANO, R. J., SNIDER, L. M. & ORLIN, M. N. 2004. Recent advances in physical and occupational therapy for children with cerebral palsy. *Seminars in Pediatric Neurology,* 11**,** 66-77.

PALLANT, J. 2007. *SPSS survival manual: a step by step guide to data analysis using SPSS for Windows,* Maidenhead, Open University Press.

PALSBO, S. E. & HOOD-SZIVEK, P. 2012. Effect of robotic-assisted three-dimensional repetitive motion to improve hand motor function and control in children with handwriting deficits: a nonrandomized phase 2 device trial. *Am J Occup Ther,* 66**,** 682-90.

PANETH, N. 2008. Establishing the diagnosis of cerebral palsy. *Clin Obstet Gynecol,* 51**,** 742-8.

PERLSTEIN, M. A. 1952. Infantile cerebral palsy; classification and clinical correlations. *J Am Med Assoc,* 149**,** 30-4.

PRESTON, N., CLARKE, M. & BHAKTA, B. 2011. Development of a framework to define the functional goals and outcomes of botulinum toxin A spasticity treatment relevant to the child and family living with cerebral palsy using the International Classification of Functioning, Disability and Health for Children and Youth. *J Rehabil Med,* 43**,** 1010-5.

PRESTON, N., WEIGHTMAN, A., CULMER, P., LEVESLEY, M., BHAKTA, B. & MON-WILLIAMS, M. 2014a. The Cerebral Palsy Kinematic Assessment Tool (CPKAT): feasibility testing of a new portable tool for the objective evaluation of upper limb kinematics in children with cerebral palsy in the non-laboratory setting. *Disabil Rehabil Assist Technol***,** 1-6.

PRESTON, N., WEIGHTMAN, A., GALLAGHER, J., HOLT, R., CLARKE, M., MON-WILLIAMS, M., LEVESLEY, M. & BHAKTA, B. 2014b. Feasibility of school-based computer-assisted robotic gaming technology for

upper limb rehabilitation of children with cerebral palsy. *Disability and Rehabilitation: Assistive Technology,* 0**,** 1-8.

QIU, Q., RAMIREZ, D. A., SALEH, S., FLUET, G. G., PARIKH, H. D., KELLY, D. & ADAMOVICH, S. V. 2009. The New Jersey Institute of Technology Robot-Assisted Virtual Rehabilitation (NJIT-RAVR) system for children with cerebral palsy: a feasibility study. *J Neuroeng Rehabil,* 6**,** 40.

RANDALL, M., CARLIN, J. B., CHONDROS, P. & REDDIHOUGH, D. 2001. Reliability of the Melbourne Assessment of Unilateral Upper Limb Function. *Developmental Medicine and Child Neurology,* 43**,** 761-767.

RANDALL, M., IMMS, C. & CAREY, L. 2008. Establishing validity of a modified Melbourne assessment for children ages 2 to 4 years. *American Journal of Occupational Therapy,* 62**,** 373-383.

RANDALL, M., PALLANT, J., IMMS, C. & CAREY, L. 2010. Psychometric properties of the 'extended' Melbourne Assessment of Unilateral Upper limb Function based on Rasch analysis. *Developmental Medicine and Child Neurology. Conference: 5th Biennial Conference of the Australasian Academy of Cerebral Palsy and Developmental Medicine Christchurch New Zealand. Conference Start,* 52.

READ, J. C., MACFARLANE, S. J. & CASEY, C. Endurability, engagement and expectations: measuring children's fun. Conference on interaction design and children, 2002 Eindhoven University of Technology, Eindhoven

REDDIHOUGH, D. S. & COLLINS, K. J. 2003. The epidemiology and causes of cerebral palsy. *Aust J Physiother,* 49**,** 7-12.

RETHLEFSEN, S. A., RYAN, D. D. & KAY, R. M. 2010. Classification systems in cerebral palsy. *Orthop Clin North Am,* 41**,** 457-67.

RICKEN, A. X. C., BENNETT, S. J. & SAVELSBERGH, G. J. P. 2005. Coordination of Reaching in Children with Spastic Hemiparetic Cerebral Palsy Under Different Task Demands. *Motor Control,* 9**,** 357-371.

ROBINSON, M. N., PEAKE, L. J., DITCHFIELD, M. R., REID, S. M., LANIGAN, A. & REDDIHOUGH, D. S. 2009. Magnetic resonance imaging findings in a population-based cohort of children with cerebral palsy. *Dev Med Child Neurol,* 51**,** 39-45.

ROSENBAUM, P., PANETH, N., LEVITON, A., GOLDSTEIN, M., BAX, M., DAMIANO, D., DAN, B. & JACOBSSON, B. 2007. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol Suppl,* 109**,** 8-14.

ROSENBAUM, P., RUSSELL, D., CADMAN, D., GOWLAND, C., JARVIS, S. & HARDY, S. 1990. Issues in measuring change in motor function in children with cerebral palsy: a special communication. *Phys Ther.,* 70**,** 125-31.

ROSENBAUM, P. & STEWART, D. 2004. The World Health Organization International Classification of Functioning, Disability, and Health: a model to guide clinical thinking, practice and research in the field of cerebral palsy. *Semin Pediatr Neurol,* 11**,** 5-10.

RUSSO, R. N., CROTTY, M., MILLER, M. D., MURCHLAND, S., FLETT, P. & HAAN, E. 2007. Upper Limb Botulinum toxin A injection and occupational therapy in children with hemiplegic cerebral palsy

identified from a Population register: A singleblind randomised controlled trial. *Pediatrics,* 119**,** 1149–1158.

SAKZEWSKI, L. 2012. Bimanual therapy and constraint-induced movement therapy are equally effective in improving hand function in children with congenital hemiplegia. *Journal of Physiotherapy,* 58**,** 59.

SAKZEWSKI, L., BOYD, R. & ZIVIANI, J. 2007. Clinimetric properties of participation measures for 5- to 13-year-old children with cerebral palsy: A systematic review. *Developmental Medicine and Child Neurology,* 49**,** 232-240.

SAKZEWSKI, L., ZIVIANI, J., ABBOTT, D. F., MACDONELL, R. A., JACKSON, G. D. & BOYD, R. N. 2011a. Participation Outcomes in a Randomized Trial of 2 Models of Upper-Limb Rehabilitation for Children With Congenital Hemiplegia. *Archives of Physical Medicine and Rehabilitation,* 92**,** 531-539.

SAKZEWSKI, L., ZIVIANI, J., ABBOTT, D. F., MACDONELL, R. A., JACKSON, G. D. & BOYD, R. N. 2011b. Randomized trial of constraint-induced movement therapy and bimanual training on activity outcomes for children with congenital hemiplegia. *Dev Med Child Neurol,* 53**,** 313-20.

SAKZEWSKI, L., ZIVIANI, J. & BOYD, R. 2009. Systematic review and meta-analysis of therapeutic management of upper-limb dysfunction in children with congenital hemiplegia. *Pediatrics,* 123**,** e1111-22.

SALATI, R., BORGATTI, R., GIAMMARI, G. & JACOBSON, L. 2002 Oculomotor dysfunction in cerebral visual impairment following perinatal hypoxia. *Dev Med Child Neurol,* 44**,** 542-550.

SAMILSON, R. L. (ed.) 1975. *Orthopaedic aspects of cerebral palsy,* London: Spastics International Medical Publications [in association with] Heinemann Medical Books; Philadelphia, Lippincott.

SANDERS, J., MCCONNELL, S., KING, R., LANFORD, A., MONTPETIT, K., GATES, P., RICH, M., SHEPHERD, K., CUPP, T., HAYNES, R., BUSH, P., TAHIR, F., SANTIAGO, J., LIGHTER, D., SMRCINA, C., NIEDERPRUEM, M., MCDONALD, C. & CURRY, D. 2006. A prospective evaluation of the WeeFIM in patients with cerebral palsy undergoing orthopaedic surgery. *Journal of Pediatric Orthopaedics,* 26**,** 542-546.

SANDLUND, M., MCDONOUGH, S. & HAGER-ROSS, C. 2009. Interactive computer play in rehabilitation of children with sensorimotor disorders: a systematic review. *Dev Med Child Neurol,* 51**,** 173-9.

SCHNEIBERG, S., MCKINLEY, P. A., SVEISTRUP, H., GISEL, E., MAYO, N. E. & LEVIN, M. F. 2010. The effectiveness of task-oriented intervention and trunk restraint on upper limb movement quality in children with cerebral palsy. *Developmental Medicine & Child Neurology,* 52**,** e245-e253.

SCHULZ, K. F., ALTMAN, D. G. & MOHER, D. 2010. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ,* 340.

SCHULZ, K. F. & GRIMES, D. A. 2002. Blinding in randomised trials: hiding who got what. *Lancet,* 359**,** 696-700.

SCOPE. 2013. *Disability facts and figures* [Online]. Available: http://www.scope.org.uk/news/disability-2012/disability-facts-and-figures.

SCOTT, N. W., MCPHERSON, G. C., RAMSAY, C. R. & CAMPBELL, M. K. 2002. The method of minimization for allocation to clinical trials. a review. *Control Clin Trials,* 23**,** 662-74.

SCRUTTON, D. 2004. Introduction. *In:* SCRUTTON, D., DAMIANO, D. L. & MAYSTON, M. (eds.) *Management of the motor disorders of children with cerebral palsy.* 2 ed. Cambridge: Cambridge University Press.

SEOK, P. M., YOUB, C. C., MIN, L. K., HYUK, S. K., CHOI, I., CHO, T., YOO, W., LEE, S., KWON, D. & KIM, T. 2012. Rasch analysis of the pediatric outcomes data collection instrument in 720 patients with cerebral palsy. *J Pediatr Orthop.,* 32**,** 423-31. doi 10.1097/BPO.0b013e31824b2a1f.

SHEVELL, M. I., MAJNEMER, A. & MORIN, I. 2003. Etiologic yield of cerebral palsy: a contemporary case series. *Pediatr Neurol,* 28**,** 352-9.

SKOLD, A., HERMANSSON, L., KRUMLINDE-SUNDHOLM, L., ELIASSON, A. & C 2011. Development and evidence of validity for the Children's Hand-use Experience Questionnaire (CHEQ). *Developmental Medicine and Child Neurology,* 53**,** 436-442.

SKOLD, A., KRUMLINDE-SUNDHOLM, L., HERMANSSON, L., ELIASSON, A. & C 2009. Development of children's hand-use experience questionnaire - CHEQ. *Developmental Medicine and Child Neurology. Conference: 21st Annual Meeting of the European Academy of Childhood Disability Vilnius Lithuania. Conference Start,* 51.

SMITH, V., DEVANE, D., BEGLEY, C. M., CLARKE, M. & HIGGINS, S. 2009. A systematic review and quality assessment of systematic reviews of randomised trials of interventions for preventing and treating preterm birth. *European Journal of Obstetrics & Gynecology and Reproductive Biology,* 142**,** 3-11.

SMITH, Y., HONG, E. & PRESSON, C. 2000. Normative and validation studies of the Nine-hole Peg Test with children. *Perceptual & Motor Skills,* 90**,** 823-43.

SMITHERMAN, J., DAVIDS, J., TANNER, S., HARDIN, J., WAGNER, L., PEACE, L. & GIDEWALL, M. 2011. Functional outcomes following single-event multilevel surgery of the upper extremity for children with hemiplegic cerebral palsy. *J Bone Joint Surg Am.,* 93**,** 655-61. doi 10.2106/JBJS.J.00295.

SORSDAHL, A., MOE-NILSSEN, R. & STRAND, L. 2008. Observer reliability of the gross motor performance measure and the quality of upper extremity skills test, based on video recordings. *Developmental Medicine and Child Neurology,* 50**,** 146-151.

SPETH, L., LEFFERS, P., JANSSEN-POTTEN, Y. & VLES, J. 2005. Botulinum toxin A and upper limb functional skills in hemiparetic cerebral palsy: A randomized trial in children receiving intensive therapy. *Developmental Medicine and Child Neurology,* 47**,** 468-473.

SRŠEN, K. 2012. Evaluation measures for children with cerebral palsy. *Eastern Journal Medicine,* 17**,** 156-165.

STANLEY, F. J. & ALBERMAN, E. D. (eds.) 1984. *The Epidemiology of the cerebral palsies,* England: London : Spastics International Medical Publications ; Philadelphia : Lippincott, 1984.

STEENBEEK, D., KETELAAR, M., GALAMA, K. & GORTER, J. W. 2007. Goal attainment scaling in paediatric rehabilitation: A critical review of

the literature. *Developmental Medicine and Child Neurology,* 49**,** 550-556.

STERLING, C., TAUB, E., DAVIS, D., RICKARDS, T., GAUTHIER, L. V., GRIFFIN, A. & USWATTE, G. 2013. Structural Neuroplastic Change After Constraint-Induced Movement Therapy in Children With Cerebral Palsy. *Pediatrics,* 131**,** e1664-e1669.

STRAUSS, D., CABLE, W. & SHAVELLE, R. 1999. Causes of excess mortality in cerebral palsy. *Dev Med Child Neurol,* 41**,** 580-5.

STRAUSS, D. J., SHAVELLE, R. M. & ANDERSON, T. W. 1998. Life expectancy of children with cerebral palsy. *Pediatr Neurol,* 18**,** 143-9.

STREINER, D. L. & NORMAN, G. R. 2003. *Health Measurement Scales,* Oxford, Oxford University Press.

STUCKI, G., DALTROY, L., KATZ, J., JOHANNESSON, M. & LIANG, M. 1996. Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. *J Clin Epidemiol.,* 49**,** 711-7.

STUCKI, G. & SIGL, T. 2003. Assessment of the impact of disease on the individual. *Best Pract Res Clin Rheumatol,* 17**,** 451-73.

SURVEILLANCE OF CEREBRAL PALSY IN EUROPE 2000. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. Surveillance of Cerebral Palsy in Europe (SCPE). *Dev Med Child Neurol,* 42**,** 816-24.

SUTCLIFFE, T. L., GAETZ, W. C., LOGAN, W. J., CHEYNE, D. O. & FEHLINGS, D. L. 2007. Cortical reorganization after modified constraint-induced movement therapy in pediatric hemiplegic cerebral palsy. *J Child Neurol,* 22**,** 1281-7.

SVENSSON, E. 2001. Guidelines to statistical evaluation of data from rating scales and questionnaires. *J Rehabil Med,* 33**,** 47-8.

TAUB, E. 2004. Harnessing brain plasticity through behavioral techniques to produce new treatments in neurorehabilitation. *Am Psychol,* 59**,** 692-704.

TAUB, E., CRAGO, J. E. & USWATTE, G. 1998. Constraint-induced movement therapy: A new approach to treatment in physical rehabilitation. *Rehabilitation Psychology,* 43**,** 152-170.

TAUB, E. & USWATTE, G. 2003. Constraint-induced movement therapy: bridging from the primate laboratory to the stroke rehabilitation laboratory. *J Rehabil Med***,** 34-40.

TAUB, E., USWATTER, G. & PIDIKITI, R. 1999. Constraint-Induced Movement Therapy: A New Family of Techniques with Broad Application to Physical Rehabilitation - A Clinical Review. *Journal of Rehabilitation Research and Development,* 36.

TENNANT, A. 2007. Goal attainment scaling: current methodological challenges. *Disabil Rehabil.,* 29**,** 1583-8.

TENNANT, A. & CONAGHAN, P. G. 2007. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research,* 57**,** 1358-1362.

TENNANT, A., MCKENNA, S. P. & HAGELL, P. 2004. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health,* 7 Suppl 1**,** S22-6.

THOMPSON, A. J., JARRETT, L., LOCKLEY, L., MARSDEN, J. & STEVENSON, V. L. 2005. Clinical management of spasticity. *J Neurol Neurosurg Psychiatry,* 76**,** 459-63.

THORLEY, M., LANNIN, N., CUSICK, A., NOVAK, I. & BOYD, R. 2012a. Construct validity of the Quality of Upper Extremity Skills Test for children with cerebral palsy. *Developmental Medicine and Child Neurology,* 54**,** 1037-1043.

THORLEY, M., LANNIN, N., CUSICK, A., NOVAK, I. & BOYD, R. 2012b. Reliability of the quality of upper extremity skills test for children with cerebral palsy aged 2 to 12 years. *Phys Occup Ther Pediatr.,* 32**,** 4-21. doi 10.3109/01942638.2011.602389. Epub 2011 Aug 15.

UNITED NATIONS 2006. Annex 1, A/61/611 Final report of the Ad Hoc Committee on a Comprehensive and Integral International Convention on the Protection and Promotion of the Rights and Dignity of Persons with Disabilities.

VAN HEDEL, H. & WICK, K. 2011. Validity and reliability of easy-to-apply hand function and capacity tests in children with cerebral palsy. *Neurorehabilitation and Neural Repair. Conference: 1st European NeuroRehabilitation Congress, ENRC,* 26**,** 403-404.

VENETSANOU, F. K., A. AGGELOUSSIS, N. SERBEZIS, V. TAXILDARIS, K. 2007. Use of the Bruininks–Oseretsky Test of Motor Proficiency for identifying children with motor impairment. *Developmental Medicine & Child Neurology,* 49**,** 846-848.

VERKERK, G. J. Q., WOLF, M. J. M. A. G., LOUWERS, A. M., MEESTER-DELVER, A. & NOLLET, F. 2006. The reproducibility and validity of the Canadian Occupational Performance Measure in parents of children with disabilities. *Clinical Rehabilitation,* 20**,** 980-988.

VOERMAN, G. E., GREGORIC, M. & HERMENS, H. J. 2005. Neurophysiological methods for the assessment of spasticity: the Hoffmann reflex, the tendon reflex, and the stretch reflex. *Disabil Rehabil,* 27**,** 33-68.

VOS-VROMANS, D. C. W. M., KETELAAR, M. & GORTER, J. W. 2005. Responsiveness of evaluative measures for children with cerebral palsy: The Gross Motor Function Measure and the Pediatric Evaluation of Disability Inventory. *Disability and Rehabilitation,* 27**,** 1245-1252.

WAGNER, L. & DAVIDS, J. 2012. Assessment tools and classification systems used for the upper extremity in children with cerebral palsy. *Clin Orthop Relat Res.,* 470**,** 1257-71. doi 10.1007/s11999-011-2065-x.

WALLEN, M., BUNDY, A., PONT, K. & ZIVIANI, J. 2009a. The revised pediatric motor activity log to measure upper limb outcome in children with hemiplegic cerebral palsy. *Developmental Medicine and Child Neurology. Conference: 3rd International Cerebral Palsy Conference Sydney, NSW Australia. Conference Start,* 51.

WALLEN, M., BUNDY, A., POT, K. & ZIVIAN, J. 2009b. Psychometric properties of the Pediatric Motor Activity Log used for children with cerebral palsy. *Developmental Medicine and Child Neurology,* 51**,** 200-208.

WALLEN, M., O'FLAHERTY, S. J. & WAUGH, M. C. 2007. Functional outcomes of intramuscular botulinum toxin type a and occupational therapy in the upper limbs of children with cerebral palsy: a randomized controlled trial. *Arch Phys Med Rehabil,* 88**,** 1-10.

WALLEN, M., ZIVIANI, J., NAYLOR, O., EVANS, R., NOVAK, I. & HERBERT, R. D. 2011. Modified constraint-induced therapy for children with hemiplegic cerebral palsy: a randomized trial. *Dev Med Child Neurol,* 53**,** 1091-9.

WANG, H., LIAO, H. & HSIEH, C. 2006. Reliability, sensitivity to change, and responsiveness of the peabody developmental motor scales-second edition for children with cerebral palsy. *Physical Therapy,* 86**,** 1351-9.

WANG, M. & REID, D. 2011. Virtual reality in pediatric neurorehabilitation: attention deficit hyperactivity disorder, autism and cerebral palsy. *Neuroepidemiology,* 36**,** 2-18.

WEIGHTMAN, A., PRESTON, N., LEVESLEY, M., HOLT, R., MON-WILLIAMS, M., CLARKE, M., COZENS, A. J. & BHAKTA, B. 2011. Home based computer-assisted upper limb exercise for young children with cerebral palsy: a feasibility study investigating impact on motor control and functional outcome. *J Rehabil Med,* 43**,** 359-63.

WEIGHTMAN, A. P. H., PRESTON, N., HOLT, R., ALLSOP, M., LEVESLEY, M. & BHAKTA, B. 2009. Engaging children in healthcare technology design: developing rehabilitation technology for children with cerebral palsy. *Journal of Engineering Design,* 21**,** 579-600.

WILSON, B. K., BJ. CRAWFORD, SG. 2000. Interrater reliability of the Bruininks-Oseretsky Test of Motor Proficiency-long form. *Adapted physical activity quarterly,* 17**,** 95.

WILSON, M. 2005. *Constructing Measures: An Item Response Modeling Approach,* New Jersey, Lawrence Erlbaum Associates Inc.

WINKELS, D. G., KOTTINK, A. I., TEMMINK, R. A., NIJLANT, J. M. & BUURKE, J. H. 2013. Wii-habilitation of upper extremity function in children with cerebral palsy. An explorative study. *Dev Neurorehabil,* 16**,** 44-51.

WISANSKOONWONG, P., FAHY, K. & HASTIE, C. 2011. The effectiveness of medical interventions aimed at preventing preterm birth: A literature review. *Women and Birth,* 24**,** 141-147.

WORLD HEALTH ORGANISATION 2001. Towards a Common Language for Functioning, Disability and Health: ICF. Geneva, Switzerland: World Health Organization.

YANG, T. F., FU, C. P., KAO, N. T., CHAN, R. C. & CHEN, S. J. 2003. Effect of botulinum toxin type A on cerebral palsy with upper limb spasticity. *Am J Phys Med Rehabil,* 82**,** 284-9.

ZAUNER, A. & MUIZELAAR, J. P. 1997. Brain metabolism and cerebral blood flow. *In:* REILLY, P. & BULLOCK, R. (eds.) *Head Injury.* London: Chapman & Hall.