Designing and evaluating technologies for virtual reality therapies that promote neuroplasticity

by

Danica Kyryllo

A thesis submitted in conformity with the requirements for the degree of Master of Health Science
Institute of Biomaterials and Biomedical Engineering
University of Toronto

© Copyright by Danica Kyryllo 2014
Designing and evaluating technologies for virtual reality therapies that promote neuroplasticity

Danica Kyryllo

Master of Health Science

Institute of Biomaterials and Biomedical Engineering
University of Toronto

2014

Abstract

Increasingly, virtual reality therapy (VRT) technologies are being used to augment pediatric rehabilitation. The mechanisms underlying success/failure of VRTs are not well understood. This thesis proposed an innovative 3-phase framework for evaluating VRT technologies with respect to neuroplasticity based on results of a scoping review of 21 studies. A case study was undertaken to demonstrate use of the framework to design and evaluate ‘Musical Steps’, a VRT technology aimed at promoting heel contact in toe-walking children. 5 therapists and 4 children were engaged in this study. The system accurately detected 88%(SD=7%) of heel contacts and was rated positively in usability testing (phase 1). Feasibility studies indicated that, while enjoyable, children did not understand the feedback provided and hence, heel contact was not increased (phase 2). These findings will direct future reiterations prior to evaluating clinical impact (phase 3). The proposed framework may enhance design and translation of therapeutically relevant VRTs.
Acknowledgments

First, I would like to thank my husband, Justin, for always believing in me and providing me with the love and support I needed to persevere. Your unwavering optimism and encouragement has inspired me to get through it all. Thank you.

To my incredible thesis supervisor, Dr. Elaine Biddiss, I am so grateful for your enthusiasm, support, and guidance through thick and thin. Thank you so much for giving me this wonderful opportunity to work with you.

To Ajmal Khan, I am indebted to you for all your help and stimulating conversations. Thank you for your charitable nature and answering my seemingly endless questions. I could not have got to this point without you, so thank you.

I would like to thank my committee members, Dr. Virginia Wright and Dr. Karl Zabjek for their guidance and thought-provoking discussions throughout my project leading to this exciting end result.

I would also like to thank Claire Mackenzie for all of her help and support in recruiting participants.

I would like to thank the members of the PRISM lab for helping me through technical difficulties and answering my abundance of questions. I would particularly like to thank Nayana for coming to my rescue.

A very special thanks to all the children who participated in this study.

This research was funded in part by the Holland Bloorview Kids Rehabilitation Hospital Foundation. I am thankful to have had their support.

I would also like to acknowledge the members of the Hypertonia clinic at Holland Bloorview Kids Rehabilitation Hospital for all their efforts in trying to recruit participants.

Last, but certainly not least, to my parents, I thank you for your never-ending, unquestioning love and support. Thank you for your inspiration and motivation in times of need.
# Table of Contents

Acknowledgments ........................................................................................................ iii  
Table of Contents ........................................................................................................ iv  
List of Tables ................................................................................................................ vii  
List of Figures ............................................................................................................... ix  
Chapter 1 ......................................................................................................................... 1  
   1 Introduction ............................................................................................................... 1  
      1.1 Motivation .......................................................................................................... 1  
      1.2 Research Questions and Objectives ................................................................. 2  
      1.3 Thesis Layout .................................................................................................... 2  
Chapter 2 ......................................................................................................................... 4  
   2 Tools for Evaluating Virtual Reality Therapy for Eliciting Neuroplasticity in Children ............................................................................................................... 4  
      2.1 Introduction ...................................................................................................... 4  
      2.2 Methods ............................................................................................................ 6  
         2.2.1 Data Sources ............................................................................................... 6  
         2.2.2 Study Selection ......................................................................................... 7  
         2.2.3 Data Extraction ......................................................................................... 7  
         2.2.4 Data Analysis ............................................................................................ 8  
      2.3 Results ............................................................................................................... 8  
         2.3.1 Motivation .................................................................................................. 11  
         2.3.2 Feedback ................................................................................................... 11  
         2.3.3 Practice ..................................................................................................... 11  
         2.3.4 Complexity/Difficulty ............................................................................. 12
2.3.5 System Usability ............................................................12

2.4 Discussion .......................................................................24
  2.4.1 Key Findings ..............................................................24
  2.4.2 Recommendations: Towards a Standardized Toolkit ..........25
  2.4.3 Recommendations towards a Standardized Framework ..........27
  2.4.4 Limitations and Merits ................................................32

2.5 Conclusion .......................................................................33

Chapter 3 ...........................................................................34

3 Using the Musical Steps augment-reality system toward heel contact in toe-walking children: A case study .................................................................34

  3.1 Introduction ....................................................................34
    3.1.1 Motivation ...............................................................34
    3.1.2 Therapeutic Goal ......................................................34
    3.1.3 Physical Therapy for Toe-Walking Children .................35
    3.1.4 Auditory Biofeedback ..............................................36
    3.1.5 Gait Analysis Technologies ....................................37
    3.1.6 Study Objectives .....................................................38

  3.2 Methods ........................................................................39
    3.2.1 Musical Steps – System Design ................................39
    3.2.2 Evaluation: Application of Novel VRT Framework ..........43
    3.2.3 Data Analysis ..........................................................55

  3.3 Results ..........................................................................56
    3.3.1 Participants ............................................................56
    3.3.2 Phase 1: System Design and Usability Testing ...............58
    3.3.3 Phase 2: Feasibility Studies .....................................62
### Chapter 3: Discussion

- **3.4 Discussion** ........................................................................................................... 65
  - **3.4.1 Key Findings** .................................................................................................. 65
  - **3.4.2 Study Strengths & Limitations** ................................................................. 67
  - **3.4.3 Future Work** ................................................................................................. 73
  - **3.4.4 Clinical Implications** ................................................................................ 78

### Chapter 4: Conclusion

- **4.1 Overview** .......................................................................................................... 79
- **4.2 Summary of Contributions** ............................................................................. 80
- **4.3 Directions for Future Work** ............................................................................ 81
- **4.4 Closing Remarks** ............................................................................................. 82

### References

........................................................................................................................................... 84

### Appendices

........................................................................................................................................... 96
List of Tables

Table 1: Working definitions for the key factors affecting of neuroplasticity .......................... 9

Table 2: Motivation measures analysis.................................................................................. 13

Table 3: Reward/Feedback measures analysis........................................................................ 18

Table 4: Practice measures analysis...................................................................................... 19

Table 5: Complexity/Difficulty measures analysis................................................................. 21

Table 6: System Usability measures analysis......................................................................... 23

Table 7: Summary of measures and tools used to evaluate the system design and usability testing.......................................................................................................................................................... 50

Table 8: Summary of measures and tools used to evaluate user interaction with the Musical Steps system................................................................................................................................................. 50

Table 9: Summary of measures and tools used to evaluate motivation.................................... 52

Table 10: Summary of measures and tools used to evaluate feedback.................................... 53

Table 11: Summary of measures and tools used to evaluate practice..................................... 53

Table 12: Summary of measures and tools used to evaluate difficulty .................................... 54

Table 13: Quasi-Confusion Matrix for Heel Step Detection Algorithm Evaluation ............. 55

Table 14: Child Participant Demographics. (CP=Cerebral Palsy; MCA=Middle Cerebral Artery; ASD=Autism Spectrum Disorder; P1, P2, P3, P4 represent the first, second, third, and fourth child participants, respectively).................................................................................................................. 57

Table 15: Summary of results for user preference evaluation tools ......................................... 59

Table 16: Summary of results for usability evaluation tools .................................................... 59
Table 17: Summary of results for enjoyment/fun evaluation tools ................................. 62

Table 18: Summary of results for distraction evaluation tools ........................................ 62

Table 19: Summary of results for engagement evaluation tools ...................................... 63

Table 20: Summary of results for sense of agency evaluation tools .............................. 63

Table 21: Summary of results for achievement of therapeutic goal evaluation tools ....... 64

Table 22: Summary of results for exertion evaluation tools ......................................... 64

Table 23: Summary of results for fatigue evaluation tools ............................................. 65

Table 24: Summary of system usability issues identified by the researcher ..................... 71
List of Figures

Figure 1: Summary of technologies used in the reviewed papers .........................................................9

Figure 2: Summary of diagnoses included in the reviewed studies (CP=Cerebral Palsy; ASD=Autism Spectrum Disorder; Various Other Disorders included: Acquired Brain Injury, Acquired Muscle Dystrophy, Developmental Co-ordination Disorder, Down Syndrome, Mild Mental Retardation, Spinal Cord Injury, Severe Motor Impairment, Multiple Sclerosis, Gross Developmental Delay, Gross Motor Delay, Meningoencephalitis, and Neurofibromatosis)..10

Figure 3: Number of studies per year ..................................................................................................................10

Figure 4: Current paradigm for evaluating VRT technologies ........................................................................27

Figure 5: Proposed iterative 3-stage VRT evaluation framework .................................................................29

Figure 6: Iterative design cycle adapted from [43]. Dotted line indicates the step in the cycle where the developed framework should be applied.................................................................30

Figure 7: Previously developed iterative 3-stage VRT evaluation framework.................................38

Figure 8: Gyroscope and Accelerometer Coordinate System Orientation..................................................40

Figure 9: Shimmer 9DoF wireless sensing units were equipped with two FSRs ..................40

Figure 10: Algorithm A logic flow diagram using only FSR data to detect heel contact events ........................................................................................................................................41

Figure 11: Screens and workflow of the developed prototype clinician interface .................42

Figure 12: Placement of the Shimmer sensor units on the posterior aspect of the shoe ........45

Figure 13: Approximate placement of the FSRs on the sole of the shoe........................................45

Figure 14: Testing environment for the gait sessions was a hallway with a 10-meter runway marked on the floor ........................................................................................................................................46
Figure 15: Participant demographics for Usability Testing with Therapists. (T1, T2, T3, and T4 represent the first, second, third, and fourth therapist participants, respectively).
Chapter 1

1 Introduction

1.1 Motivation

Physical and occupational therapies are important activities in the rehabilitation of children with physical disabilities. Success in these therapies is most commonly measured via improvements on functional outcomes that may be linked to changes in physical and/or cognitive capacity through strengthening, endurance building, sensory perception, and motor learning, for instance. Motor learning depends on many aspects of a child’s abilities and therapeutic environment and can be evidenced through neuroplasticity. Neuroplasticity can be defined as “the ability of the nervous system to respond to intrinsic or extrinsic stimuli by reorganizing its structure, function and connections.” [1] In the context of goal-based motor rehabilitation, neuroplasticity relies on 4 key factors: motivation, feedback, practice, and task complexity/difficulty. [1] [2] [3] These factors are not mutually exclusive; for example, an increase in motivation or enjoyment of an activity may promote a child’s practice of a therapeutic goal. Alternatively, if a child does not feel motivated or gets bored of performing a task repetitively, as is common in therapies with children, their potential for motor learning decreases. Optimizing therapy environments for motor learning may lead to a number of important outcomes including increased patient and family satisfaction, greater functional improvements, faster recovery times and associated cost savings to the healthcare system.

In recent years, interest in virtual reality therapy (VRT) technologies as potential tools for promoting motor learning in therapy activities has increased. [4] A VRT technology can be broadly defined as a system that provides an augmented environment (e.g. through graphics, sounds, gamification) with varying levels of immersion (e.g. fully- or non-immersive environments) that is intended to enhance engagement during physical or occupational therapy. A VRT intervention is the therapeutic application of a VRT technology or system. The main difference between the technology versus the intervention is that one may investigate the technology or system independently from its application in a therapeutic setting, whereas an intervention is investigated with respect to a specific therapeutic context in which the
technology is being used. Many studies have investigated the potential of VRT technologies with respect to their impact on functional outcomes [4] [5] [6] [7] [8], but the reasons why these technologies succeed or fail in eliciting functional changes is not always well understood. In evaluating these technologies, it is important to understand what the “active ingredients” [9] of a VRT technology are in order to optimize their ability to promote factors of motor learning in children. Given the dramatic increase in interest in VRT technologies for therapeutic use in recent years, it is important that we consider how and if these technologies are being evaluated in ways that will facilitate: (1) the development of improved VRT technologies, (2) our understanding of the mechanisms by which VRT technologies elicit functional changes, and (3) the successful translation of these technologies to the clinic and home environments.

1.2 Research Questions and Objectives

Toward the development of a standardized methodology of evaluating VRT technologies, this research thesis aimed at answering two guiding questions.

Question 1: How should VRT technologies be evaluated?

- Objective 1a: Review evaluation tools used in the VRT literature
- Objective 1b: Develop a new framework for evaluating VRT technologies

Question 2: How can this framework be applied to a new VRT intervention?

- Objective 2a: Design ‘Musical Steps’, a VRT technology to address “toe-walking” in gait therapy
- Objective 2b: Evaluate ‘Musical Steps’, applying the developed framework

1.3 Thesis Layout

This thesis is presented as a compilation of two studies, which together provide a foundation towards a standardized framework for evaluating VRT technologies with respect to factors of neuroplasticity. The research is of interest to game developers, researchers and clinicians working in the area of VRT from the early usability testing phase of the design process through to the clinical evaluation of VRT interventions. The chapters following this introduction are:
• Chapter 2 investigates the current state of VRT technologies in the context of motor rehabilitation in children with disabilities with respect to the tools that have previously been utilized in the literature for evaluating the technology’s effects on factors of neuroplasticity. Based on this toolkit, a novel framework was developed toward a standard methodology for VRT technology evaluation.

• Chapter 3 presents a case study example use of the developed framework through the development and evaluation of a novel VRT technology, ‘Musical Steps’, aimed at improving heel contact in toe-walking children during gait training therapy sessions.

• Chapter 4 restates and summarizes the contributions provided by this research thesis and provides recommendations for directions of future work.
Chapter 2

2 Tools for Evaluating Virtual Reality Therapy for Eliciting Neuroplasticity in Children

2.1 Introduction

Increasingly, technologies such as virtual reality (VR), augmented reality (AR), interactive computer play (ICP) and robotics are being used to augment traditional pediatric rehabilitation therapies. A Virtual Reality Therapy (VRT) technology is broadly defined here as a system that provides an augmented environment (e.g. through graphics, sounds, gamification) with varying levels of immersion (e.g. fully- or non-immersive environments) that is intended to enhance engagement during physical or occupational therapy.

Interest in these technologies is linked to their potential for enhancing motivation and providing feedback/reward during repeated practice of a motor task. Additionally, these technologies often offer the ability to adjust the complexity or difficulty of the task to the abilities of each individual based on objective measures and training statistics. These features are fundamental to the promotion of neuroplasticity and to the success of motor learning programs. While motivation, reward, and feedback are highly interconnected, in this paper we define motivation as the impetus or drive to perform a task, which for children is very often associated with their enjoyment of an activity [10]. Reward and feedback, on the other hand, are linked to the experiences and interpretation of information resulting from performing the task [11] and must be directly in line with the therapeutic goals of the motor activity. Levac et al. [12] describe practice as one of the most relevant motor learning strategies in functionally based interventions. Practice is related to the quantity (e.g. time spent practicing, number of repetitions), structure, and schedule of performing skills or tasks. Lastly, task complexity relates to the difficulty of the task and is highly dependent on the capacity required to perform a task, which may be cognitive [13] or physical [2]. Task complexity must be continually managed to balance challenge and frustration.

A number of excellent systematic reviews have presented evidence to support the use of various virtual reality therapy (VRT) technologies in motor rehabilitation with children. [14]
These reviews present 92 studies that have proposed and investigated a large range of commercially available and custom-designed VRT games for the purposes of gross motor (e.g. reaching), fine motor (e.g. wrist flexion-extension) and sensory-motor therapies. The goal of this scoping review is not to describe the efficacy of VRTs and ICP which has been done very well in these recent reviews, but to investigate the measurement tools used in their evaluation within the framework of neuroplasticity. Specifically, our overall goal is to explore the question: How should VRT technologies be evaluated? To this purpose, our main objectives are to: (1) Review the evaluation tools used in VRT technology literature (excluding functional outcomes), and (2) Develop a new framework for evaluating VRT technologies. The review aims to discover what measures have been utilized to evaluate VRT technologies with respect to the four key factors in neuroplasticity (i.e. motivation, practice, feedback, and task complexity) and also to determine what additional measures are considered important in the design and evaluation of VRT technologies. Based on this review, we will make recommendations towards the development of a standard framework for the evaluation of VRT technologies with respect to their feasibility in eliciting motor learning and neuroplasticity. We will identify gaps where additional use or development of evaluation measures may be needed. The results of this review are applicable to game developers, researchers and clinicians working in the area of VRT from the early usability testing phase of the design process through to the clinical evaluation of VRT interventions. Specifically, this review will provide information useful for:

- Planning usability testing and pilot studies to identify areas for improvement in the design of VRT technologies prior to large-scale clinical trials and deployment
- Designing clinical trials and selecting appropriate evaluation measures
- Understanding and interpreting positive or negative results of intervention studies with respect to key concepts of neuroplasticity (e.g. if a system performed poorly, was this outcome linked to lack of appropriate feedback, insufficient motivation, inappropriate difficulty level, lack of use/repetition?) in order to provide direction for future developments
- Developing a framework that supports more standard evaluation measures and approaches that enable comparisons across studies and technologies.
2.2 Methods

2.2.1 Data Sources

In the past 10 years, 6 systematic reviews have been conducted on VRTs and ICP; 5 of these have focused on its use for children with cerebral palsy, acquired brain injury, autism spectrum disorders, attention deficit disorder and other neurological and sensorimotor disorders. [14] [5] [7] [6] [4] [8]

The studies reported in these reviews formed the basis for our investigation. Additionally, a scoping review was performed in October 2013 using computer-assisted literature searches via Scopus, ISI Web of Knowledge, PubMed, PsycINFO, EMBASE, CINAHL, Google Scholar, and Medline to find any additional studies on VRT technologies used in rehabilitation therapies published after December 2012, the date of the most recent systematic review [4]. For the scoping review, we utilized the Arskey and O’Malley [3] 6-stage framework. We deviated from this methodology slightly in that iteration was not performed in the interest of time efficiency. Additionally, the optional 6th stage of consultation with experts was not performed. The search was conducted using a combination of key words and subject headings in the following categories:

1. Technology: technology, virtual reality, robot, augmented reality, computer
2. Goal: physical therapy, occupational therapy, function, functional outcomes, neuroplasticity
4. Target demographic: child, children, cerebral palsy, autism
5. Outcomes:
   a. Motivation: motivation, engagement, play, adherence, attention, concentration, awareness, interest, enjoyment, fun, pleasure, self-efficacy, satisfaction, amusement, joy, excitement, entertainment, happiness,
   b. Complexity: difficulty, intricacy, complexity, perceived exertion, exertion
   c. Feedback/Reward: reward, feedback, usability
   d. Repetition: repetition, practice
2.2.2 Study Selection

Studies identified through the above-described search were screened with respect to the following inclusion and exclusion criteria: (1) English language, full-text, peer-reviewed publications; (2) evaluated virtual or augmented reality, computer training, socially-assistive robot, or ICP technologies for use in physical or occupational therapies; (3) described evaluation measures focused on characteristics of the technology (e.g. usability) and/or user experience (e.g. enjoyment); (4) evaluation of technology was completed or proposed for children with functional disabilities; (5) child-centric (majority of participants between ages 4 to 13). Studies were excluded if they focused solely on functional outcomes (e.g. measured via functional magnetic resonance imaging or outcome measures such as the Gross Motor Function Measure) which are well-documented in previous studies, including in the systematic reviews previously mentioned [5] [7] [6] [4] [8]. Studies that focused on global outcomes (e.g. quality of life, assessment of life habits) using tools that measured children’s general behaviours and/or traits as opposed to their direct interactions with or responses to the technology were excluded. Reference lists associated with included papers were also reviewed and additional studies identified and screened for inclusion. Eligibility was assessed by 2 independent reviewers. When necessary, study eligibility was discussed until a consensus was reached.

2.2.3 Data Extraction

Eligible studies were reviewed in detail and the following data were extracted: (i) measures reported (e.g. enjoyment, attention, engagement), (ii) type of measurement tools (e.g. equipment, questionnaires), (iii) characteristics describing the measurement tools, (iii) age and diagnosis of participants targeted in the intervention (i.e. not healthy controls), and (iv) technology used to augment the therapy.
2.2.4 Data Analysis

The measures identified in the review were categorized with respect to the four cornerstones of neuroplasticity: practice, feedback, motivation, and task complexity. Two reviewers independently categorized each measure and disagreements were discussed until consensus was reached. Measures that did not fit within one of these four categories were documented and grouped where appropriate. Measures were further described with respect to specific characteristics (i.e. subjective or objective, validated). Tools were labelled as objective if they could be quantified by measurements resulting in a single, factual value (e.g. timed metrics). Otherwise, measures which relied on some degree of personal judgement, were labelled as subjective (e.g. reporting). The validity of each measurement tool was then detailed by reviewing previous literature and validation studies. A tool was considered validated only if validation studies had been performed on the population (i.e. age, diagnosis) targeted in the study. Frequency counts were conducted to describe the demographics of the studies included with respect to technologies evaluated and which elements of neuroplasticity were included in their evaluation.

2.3 Results

In the process of this review, 126 studies were identified and screened of which 21 studies met with inclusion/exclusion criteria and were included in the detailed analysis. Within these studies, 20 measures were identified and categorized with relation to the four elements of neuroplasticity previously discussed, namely: motivation, feedback, practice, and task complexity. These findings are summarized in Table 1. To provide context for the following discussion of results, the technologies investigated (Figure 1), the diagnoses of study participants (Figure 2), and the number of studies per year (Figure 3) are documented. The most frequently utilized technology to date has been the Nintendo® Wii, while the most studied diagnosis in children has been cerebral palsy (CP). The number of studies being performed to evaluate the effect of VRT technologies has steadily increased in recent years.
Table 1: Working definitions for the key factors affecting neuroplasticity

<table>
<thead>
<tr>
<th>Working Definition</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motivation</strong></td>
<td>Attention, Distraction, Engagement, Fun/Enjoyment/Perceived Enjoyment, Initiative to play, Perceived Profit/Benefit, Motivation, Participation, Playfulness/Play, Preference for Treatment, Training Participation and Attendance</td>
</tr>
<tr>
<td><strong>Feedback</strong></td>
<td>Sense of Agency</td>
</tr>
<tr>
<td><strong>Practice</strong></td>
<td>Device Usage, Exercise Repetition, Gaming Intensity</td>
</tr>
<tr>
<td><strong>Task Complexity/Difficulty</strong></td>
<td>Difficulty, Exertion, Fatigue, Level of Activity</td>
</tr>
</tbody>
</table>

Figure 1: Summary of technologies used in the reviewed papers
Figure 2: Summary of diagnoses included in the reviewed studies (CP=Cerebral Palsy; ASD=Autism Spectrum Disorder; Various Other Disorders included: Acquired Brain Injury, Acquired Muscle Dystrophy, Developmental Co-ordination Disorder, Down Syndrome, Mild Mental Retardation, Spinal Cord Injury, Severe Motor Impairment, Multiple Sclerosis, Gross Developmental Delay, Gross Motor Delay, Meningoencephalitis, and Neurofibromatosis)

Figure 3: Number of studies per year
The following sections present a description of the measures used to evaluate the efficacy of therapy-augmenting technologies that were identified in this review with respect to: motivation, feedback, practice, and task complexity/difficulty.

2.3.1 Motivation

Measurement variables related to motivation are presented in Table 2 and include: attention, distraction, engagement, fun/enjoyment, initiative to play, interest, motivation, participation, perceived profit/benefit, playfulness/play, preference for treatment, and training participation and attendance. The most frequently measured variables were fun/enjoyment (5 studies) and motivation (5 studies). The fun/enjoyment variable was measured using the following tools: the Fun Toolkit, Time Spent Smiling, Customized Questionnaires, and a visual analog scale (VAS). The motivation variable was measured using the following tools: Customized questionnaires, the Pediatric Motivation Scale (PMS), Pediatric Volitional Questionnaire (PVQ), and verbal reports from children and their families. In total, 22 tools were discovered in the literature for measuring motivation. Of these, 27% provided objective measurements and 14% were validated for the age and diagnosis of participants in the studies. The validated subjective tools discovered were the PVQ and the Test of Playfulness. The Lokomat was the only technology used that offered a valid, objective measure related to the level of active participation of the child.

2.3.2 Feedback

Only one measure was discovered (i.e. sense of agency) and 1 subjective tool (i.e. self-report) was discovered that related to feedback outcomes as presented in Table 3. There were no objective or validated tools for evaluating the extent or success with which VRT technologies provided feedback during therapy tasks.

2.3.3 Practice

Data extracted from the eligible studies as relates to measures and tools relevant to the extent of practice/repetition elicited by the VRT technologies are presented in Table 4. The measures discovered were device usage, exercise repetition, and gaming intensity. The most frequently
used measure was device usage (6 studies). Device usage was measured via the following tools: average % of days children participated in training, time spent training/playing, game-specific metrics, and unstructured interviews. In total, 7 tools were discovered in the literature for measuring repetition. Of these, 71% provided objective measurements and only 1 of the tools was validated for the age and diagnosis of participants in the studies. The number of correct movements (via the Microsoft® Kinect) was the only validated measurement tool discovered in the literature for quantifying reward/feedback.

2.3.4 Complexity/Difficulty

Measures and tools that related to task complexity/difficulty are presented in Table 5 and included: difficulty, exertion, fatigue, and level of activity. The most frequently measured variable was fatigue (3 studies). Fatigue was measured using the following tools: customized questionnaires, therapist reporting, and time to fatigue. In total, 7 tools were discovered in the literature for measuring variables pertaining to task complexity/difficulty. Of these, 29% provided objective measurements and none of the tools were validated for the age and diagnosis of participants in the studies. All of the measures reported in the literature related to physical complexity, while no measures of cognitive complexity were reported.

2.3.5 System Usability

Additional tools were identified from the literature that did not necessarily fit into one of the above categories related to factors in neuroplasticity. In many cases, these measures (e.g. system usability) may affect all aspects of neuroplasticity as opposed to a single element. The measures identified are presented in Table 6 and included: usability and user satisfaction. The most frequently measured variable was user satisfaction (3 studies). User Satisfaction was measured via the following tools: the Health Professional Usability Questionnaire and customized questionnaires. In total, 6 tools were discovered in the literature for measuring system usability. All of these tools were subjective and none were validated.
Table 2: Motivation measures analysis

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Basic Description</th>
<th>Study</th>
<th>Age</th>
<th>Diagnosis</th>
<th>Objective</th>
<th>Subjective</th>
<th>Validated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attention</strong></td>
<td>Therapist Reporting</td>
<td>Therapists reported whether there was increased or decreased attention for various simulations.</td>
<td>[15]</td>
<td>7 &amp; 10 years</td>
<td>CP (Spastic Hemiplegia)</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time to Break Attention</td>
<td></td>
<td>Therapists observed sessions and reported the child's amount of time spent as actively engaged using a checklist with guidelines for reporting breaks in attention including child complaints, desire to do something else, or intermittent participation.</td>
<td>[16]</td>
<td>5-16 years</td>
<td>CP (Hemiplegia)</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Distraction</strong></td>
<td>Blinded Observer Reports</td>
<td>Blinded observers watched session videos and completed a customized questionnaire that rated different aspects of the participant’s engagement, including the child's distraction, on a 5-point Likert scale.</td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Engagement</strong></td>
<td>% of sessions for which the full duration of training was completed</td>
<td>The training period for the intervention was a 6-week period during which participants were monitored for their engagement for the full 45-minute duration for each training session. Criteria for evaluating engagement during sessions were not reported in the study.</td>
<td>[18]</td>
<td>9-12 years</td>
<td>CP</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fun/ Enjoyment/ Perceived Enjoyment</strong></td>
<td>Fun Toolkit</td>
<td>The Smileyometer was used which is a child-reported measure of enjoyment which uses a 5-pt Likert scale, with pictorials of smiley faces varying from awful to brilliant.</td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time Spent Smiling</td>
<td></td>
<td>Blinded observers watched session videos and recorded the length of the therapy sessions and the percentage of time that the children were smiling with</td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Age Range</td>
<td>Population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------</td>
<td>---------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customized questionnaires</td>
<td>Blinded observers watched session videos and completed a customized questionnaire that rated different aspects of each session, including the child's enjoyment, on a 5-point Likert scale.</td>
<td>4-7 years</td>
<td>Various</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Therapists completed a customized questionnaire, which included their agreement rating on a 5-point Likert scale with the statement &quot;the child enjoyed the sound feedback from the musical stairs.&quot;</td>
<td>4-7 years</td>
<td>Various</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Participants reported Yes, No, or Maybe to two survey questions asking if the simulation was fun and whether or not they would want to play it again in the future.</td>
<td>7 &amp; 10 years</td>
<td>CP (Spastic Hemiplegia)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test questionnaire</td>
<td>Using a 10-pt VAS which required participants to rate 7 items: 1. Fun doing the training 2. Fun doing the training with VR 6. Like it when therapist cheers me on 7. I would like to train with VR, if I had to train in the Lokomat.</td>
<td>8-18 years</td>
<td>CP, BS-CP, MS, SCI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NOTE: study considered this as a means of rating &quot;motivation&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A parent-reported, 10 VAS was used asking, &quot;Did your child have fun?&quot;</td>
<td>7-17 years</td>
<td>CP - Spastic diplegia and hemiplegia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children were asked to complete a post-test questionnaire of dichotomous Yes/No questions. Specifically, fun was investigated through the first question: &quot;Did you have fun climbing the stairs?&quot;</td>
<td>4-7 years</td>
<td>Various</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAS</td>
<td>A child-reported, 11-point VAS was used to measure children's perceived enjoyment.</td>
<td>6.8-14.6 years</td>
<td>CP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| NOTE: study considered this as a means of rating &quot;motivation&quot; |</p>
<table>
<thead>
<tr>
<th>Initiative to play</th>
<th>Gaming diary</th>
<th>Participants were asked to complete a daily gaming diary with the assistance of their parents. Children were asked short multiple choice survey questions, including one question which monitored whether sessions were initiated by (i) the child, (ii) a parent, or (iii) a sibling, friend, or relative. The diary was quick to complete, taking under a minute daily. Initiative to play was evaluated by comparing the percentage of sessions initiated by each the child, parents, or by siblings, friends, relatives or not reported.</th>
<th>[22]</th>
<th>6y,10m - 16y,1m</th>
<th>CP</th>
<th>✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>VAS reported by children</td>
<td>Children were asked &quot;Was the exercise interesting?&quot; and reported their response on a 10-pt VAS.</td>
<td>[20]</td>
<td>7-17 years</td>
<td>CP (Spastic diplegia and hemiplegia)</td>
<td>✓</td>
</tr>
<tr>
<td>Motivation</td>
<td>Customized questionnaires</td>
<td>Blinded observers watched session videos and completed a customized questionnaire that rated different aspects of each session, including the child's motivation, on a 5-point Likert scale.</td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Children were asked to complete a post-test questionnaire of dichotomous Yes/No questions. Specifically, motivation was investigated through the fourth question: &quot;Did the musical stairs make you want to climb the stairs?&quot;</td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Therapists complete a customized questionnaire, which included their agreement rating on a 5-point Likert scale with the statement &quot;the child was more motivated on the musical stairs.&quot;</td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
</tr>
<tr>
<td>Pediatric Motivation Scale</td>
<td>The PMS includes four questions which investigate the child’s level of enjoyment and confidence in their rehabilitation. Questions are answered on a visual analogue scale, with five smiley faces representing a range from “did not enjoy at all” to “extremely enjoyed” their therapy session.</td>
<td>[23]</td>
<td>10 years</td>
<td>ABI</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[24]</td>
<td>12-14 years</td>
<td>ABI</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

| **Verbal reports from children and their families** | Children and families' reported on the following aspects of the intervention: satisfaction with the training system, difficulty with performing the daily training; difficulties persuading children to train; enthusiasm for training, initiative to play with others, motivating effect of remote contact with therapists, motivating effect of game-like design, motivating factors of children to participate, effect of training on skills in daily, effect of system on self-confidence and ability to take on new challenges | [26] | 6-13 years | CP | ✓ |

| **Participation** | Lokomat weighted force measurements | Active participation during training sessions was measured by the Lokomat by calculating weighted average forces at the hip and knee joints for | [19] | 8-18 years | CP, BS-CP, MS, SCI | ✓ | ✓ |
| **Participants’ stance and swing phases of their gait.** These biofeedback values were unitless measures; positive values represented active participation by the child; negative values indicated that the Lokomat was carrying the child's load. |  |  |
| **Time spent on task out of total session time** | Sessions were 60 minutes in length. The amount of time children spent on task during each session was recorded by the computer system. Criteria for evaluating if children were on task or not were not reported in the study. | [16] | 5-16 years | Hemiplegia secondary to CP | ✓ |
| **Perceived Profit/Benefit** | Customized questionnaire Post-test questionnaire using a 10-pt VAS which required participants to rate 7 items: 3. Profit from training with Lokomat NOTE: study considered this as a means of rating "motivation" | [19] | 8-18 years | CP, BS-CP, MS, SCI | ✓ |
| **Playfulness/Play** | Semi-structured interview Interviews were conducted with adults who knew the participants and participated in the play activity. Adults interpreted specific changes in child behaviour as reported in questionnaires and observation grids and provided general feedback on the child's interactions with the robot. | [27] | Not Reported | ASD, MMR, SMI | ✓ |
| **Test of Playfulness** | The Test of Playfulness consists of 24 items within 4 subscales (motivation, control, suspension from reality, and framing). Each item is scored on 4-point scales for extent (3=almost always, 0=rarely or never), intensity (3=highly, 0= not), and skill (3 = highly skilled, 0 = unskilled). | [28] | 8-13 years | CP |  |
| **Training Participation and Attendance** | Attendance The number of sessions that participants attended over the course of the 6-week intervention was recorded to find the percent attendance rate. | [18] | 9-12 years | CP | ✓ |
| **Time Spent Training** | The recommended training was for 30 minutes per day. The number of days where participants trained over | [26] | 6-13 years | CP | ✓ |
this amount out of the total 1260 days of training were used to calculate the percentage of days spent training over the recommended daily duration.

The aim of the project was for participants to spend at least 30 minutes training per day during the intervention period of 140 days. The average time spent training per day was calculated and compared to the recommended amount.

The aim of the project was for participants to spend a total of 70 hours in training during the intervention period of 140 days. The total average number of hours spent training was calculated and compared to the recommended amount.

**Preference for Treatment**

- **Customized questionnaires**
  - An ad hoc satisfaction questionnaire was used to gather children's feedback on 7 items regarding their preference for the treatment.
  - Post-test questionnaire using a 10-pt VAS which required participants to rate 7 items:
    - 7. I would like to train with VR, if I had to train in the Lokomat
  - NOTE: study considered this as a means of rating "motivation"

### Table 3: Reward/Feedback measures analysis

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Basic Description</th>
<th>Study</th>
<th>Age</th>
<th>Diagnosis</th>
<th>Objective</th>
<th>Subjective</th>
<th>Validated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sense of Agency</strong></td>
<td>Self-reporting</td>
<td>Participants were asked to evaluate whether or not they were responsible for producing movements displayed on the screen.</td>
<td>[30]</td>
<td>8-16 years</td>
<td>CP</td>
<td>Subjective</td>
<td>✔️</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Practice measures analysis

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Basic Description</th>
<th>Study</th>
<th>Age</th>
<th>Diagnosis</th>
<th>Objective</th>
<th>Subjective</th>
<th>Validated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Device Usage</strong></td>
<td>Average % of days children participated in training</td>
<td>The average number of days participants trained out of the 140 scheduled training days was recorded.</td>
<td>[26]</td>
<td>6-13 years</td>
<td>CP</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time spent training/playing</td>
<td>The average number of minutes children spent training daily was recorded.</td>
<td>[26]</td>
<td>6-13 years</td>
<td>CP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The total average hours children spent training over the course of the intervention was recorded.</td>
<td>[26]</td>
<td>6-13 years</td>
<td>CP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>User played games which promoted movements of the upper limb in the transverse plane using a joystick which provided some support. The total number of times the system was used over their 4-week home usage of the system was recorded by the HB-RES software.</td>
<td>[31]</td>
<td>5-16 years</td>
<td>CP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The training program in the proposed study includes 16 tasks/games, which each have their own specific measures of task completion including time spent on exercise and time per repetition.</td>
<td>[32]</td>
<td>8-18 years (proposed)</td>
<td>CP - congenital hemiplegia (proposed)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>User played games which promoted movements of the upper limb in the transverse plane using a joystick which provided some support. The total time each participant spent playing on the device over their 4-week home usage of the system was recorded by the HB-RES software.</td>
<td>[31]</td>
<td>5-16 years</td>
<td>CP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>An electrogoniometer was used to determine the average repetition hold time and the time to complete each repetition for each round of 2 minute exercise.</td>
<td>[20]</td>
<td>7-17 years</td>
<td>CP - Spastic diplegia and hemiplegia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game-specific metrics</td>
<td>The training program in the proposed study includes 16 tasks/games, which each have their own specific measures of task completion including the following: % increase of balloon, % correct (movements), time spent on exercise, # pieces missed, balance distribution, time per repetition, accuracy of object path. User played games which promoted movements of the upper limb in the transverse plane using a joystick which provided some support. The number of inward and outward movements they made over their 4-week home usage of the system was recorded by the HB-RES software.</td>
<td>[32]</td>
<td>8-18 years (proposed)</td>
<td>CP - congenital hemiplegia (proposed)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------</td>
<td>----------------------</td>
<td>------------------------------------</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstructured interview</td>
<td>Participants and their parents/guardians/teacher were interviewed to gather feedback on the estimated usage of the device and games over the course of the intervention.</td>
<td>[31]</td>
<td>5-16 years</td>
<td>CP</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise Repetition</td>
<td>Number of exercise repetitions</td>
<td>An electrogoniometer was used to determine the number of repetitions completed for each round of 2 minute exercise. Sessions were video recorded and reviewed to count the total number of stairs climbed and the number of climbed using a reciprocal pattern. The primary outcome measure was the percentage of reciprocal steps taken for sessions with and without feedback. Sessions were video recorded and reviewed to count the total number of stairs climbed and the number of climbed using a reciprocal pattern. The primary outcome measure was the percentage of reciprocal steps taken for sessions with and without feedback.</td>
<td>[20]</td>
<td>7-17 years</td>
<td>CP - Spastic diplegia and hemiplegia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of correct movements</td>
<td>The Kinect was used to measure whether or not movements performed by participants corresponded to movements required by therapists. The system automatically counted the number of correct movements.</td>
<td>[34]</td>
<td>14 years</td>
<td>CP &amp; AMD</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Gaming intensity was evaluated through the use of the gaming diary responses by taking into account number of sessions per week, and the average time spent for the sessions in each week.

<table>
<thead>
<tr>
<th>Gaming Intensity</th>
<th>Tool</th>
<th>Basic Description</th>
<th>Study</th>
<th>Age</th>
<th>Diagnosis</th>
<th>Objective</th>
<th>Subjective</th>
<th>Validated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty</td>
<td>Difficulty Level</td>
<td>Therapists adjusted the level of difficulty of the training by using increasingly difficult motor challenges (ex. more repetitions required, higher load).</td>
<td>[22]</td>
<td>6 y,10m - 16y,1m</td>
<td>CP</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exertion</td>
<td>Customized questionnaires</td>
<td>Blinded observers watched session videos and completed a customized questionnaire that rated different aspects of each session, including the child's exertion, on a 5-point Likert scale. Therapists complete a customized questionnaire, which included their agreement rating on a 5-point Likert scale with the statement &quot;the child exerted themselves more on the musical stairs.&quot; Post-test questionnaire using a 10-pt VAS which required participants to rate 7 items: 4. Effort for training with VR 5. Effort for training without VR. NOTE: considered this as a means of rating &quot;motivation&quot;</td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatigue</td>
<td>Customized questionnaires</td>
<td>Blinded observers watched session videos and completed a customized questionnaire that rated different aspects of each session, including the child's fatigue, on a 5-point Likert scale.</td>
<td>[17]</td>
<td>4-7 years</td>
<td>Various</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **Level of Activity** | Energy Expenditure | Therapists were asked to complete a post-test questionnaire of dichotomous Yes/No questions. Specifically, fatigue was investigated through the second question: "Do you feel tired?"

Therapist Reporting

Therapists reported whether or not there were demonstrations of fatigue in various simulations

| Time to Fatigue | Therapists observed sessions and reported the child's amount of time to fatigue using a checklist with guidelines for reporting fatigue including child verbal report, physical signs, discontinuance of the activity, fidgeting, decreased ability or effort to continue, or increased need for verbal cues.

| **Level of Activity** | Energy Expenditure | MET values were calculated by the SenseWear Pro3 Armband which was attached to the participant's upper arm. Time participants spent above 3 METs was considered time spent 'physically active'.

Metabolic Equivalent Units (METs) are a relative unit of a person's working to resting metabolic rate. The amount of time participants spent as 'physically active' (activity levels above 3 METs) and at 'vigorous' level of activity (6–9 METs) was measured via a SenseWear Pro3 Armband. Increase in time spent as 'physically active' was compared during gaming weeks with the baseline average time.

The SenseWear Pro3 Armband was attached to the participant's upper arm and used the sensors signals to calculate energy expenditure.

Accelerometry

The SenseWear Pro3 Armband was attached to the participant's upper arm and used biaxial accelerometers to monitor the number of steps taken.

| [17] 4-7 years | Various |
| [15] 7 & 10 years | CP (Spastic Hemiplegia) |
| [16] 5-16 years | hemiplegia secondary to CP |
| [22] 6y,10m-16y,1m | CP |
| [22] 6y,10m-16y,1m | CP |
| [22] 6y,10m-16y,1m | CP |
| [22] 6y,10m-16y,1m | CP |
Table 6: System Usability measures analysis

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Basic Description</th>
<th>Study</th>
<th>Age</th>
<th>Diagnosis</th>
<th>Objective</th>
<th>Subjective</th>
<th>Validated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usability</strong></td>
<td>Unstructured interview</td>
<td>Participants and their parents/guardians/teacher were interviewed to gather feedback on the following: use at school vs. at-home, set-up time and its effect on enthusiasm to play, range of games, impact on confidence when sitting and moving, impact on degree of independence, and opportunity for the child to interact with classmates, siblings and parents.</td>
<td>[33]</td>
<td>7y.3m – 16y.2m</td>
<td>CP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>User Satisfaction</strong></td>
<td>Health professional usability questionnaire</td>
<td>A 70-question, standard questionnaire reported by physical therapists using a variety of reporting methods including VAS, multiple choice, and open-ended questions.</td>
<td>[21]</td>
<td>6.8-14.6 years</td>
<td>CP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Customized Questionnaire</td>
<td>Children were asked to report on a 5-point scale their agreement with 26 items grouped in 5 categories for user satisfaction with the system.</td>
<td>[35]</td>
<td>7-16 years</td>
<td>CP Spastic Hemi-or Tetraplegia</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children were asked to complete a post-test questionnaire of dichotomous Yes/No questions. Specifically, user satisfaction was investigated through the last three questions: &quot;Were the sounds too loud?&quot;, &quot;Did you like the sounds?&quot;, and &quot;Were the sensors annoying to wear?&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A 24-question user satisfaction questionnaire was used based on a modified version of a previously developed questionnaire. Children reported their agreement to the questions on a 5-point Likert scale.</td>
<td>[21]</td>
<td>6.8-14.6 years</td>
<td>CP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
2.4 Discussion

2.4.1 Key Findings

The importance of measuring motivation and enjoyment is fairly well-established in the literature given the frequency (67%, n=14 studies) with which studies investigated this factor. Surprisingly, only 38% of studies (n=8) quantified indicators of repetition and practice with the VRT systems. Task complexity/difficulty (29%, n=6 studies) and experience/interpretation of feedback (5%, n=1 study) were the least consistently measured of the four factors related to neuroplasticity. While 62% of studies (n=13) reported on the effectiveness of the VRT system with respect to at least one of the four factors of neuroplasticity, only 33% touched on at least 2, 14% on 3 or more, and no study looked at all 4. Of the 37 tools discovered in the VRT literature review, only 11% (n=4) were validated, which were: (i) the Pediatric Volitional Questionnaire, (ii) Lokomat force measurements, (iii) the Test of Playfulness, and (iv) the number of correct movements (Microsoft ® Kinect). The majority of tools were subjective with only 35% (n=13) being objective.

Other evaluation measures that were reported in the studies most commonly focused on system usability and user interfaces. User interaction and system usability were evaluated in 29% (n=6) of the studies. Of note, the usability of the system may directly or indirectly influence child-system interactions and the performance of the technology with respect to eliciting motivation, providing effective feedback, managing task complexity, and encouraging practice/repetition.

The results of this study suggest that our understanding of users’ interactions with VRTs is quite limited with respect to how and if feedback is effectively interpreted, the structure of practice that is achieved with respect to frequency, duration, etc. of targeted therapy tasks, and if the level of complexity/difficulty of tasks within the game context is appropriate and optimally managed. Without a better understanding of these elements, it is not possible to comment on the source of functional improvements if/when observed. These functional improvements may be achieved through processes of motor learning, but they may also be related to processes completely outside of motor learning e.g. strength building, endurance etc.
As such, functional improvements realized in VRT interventions are often outputs of a “black box”.

2.4.2 Recommendations: Towards a Standardized Toolkit

i. **Conduct more comprehensive evaluations.** To understand the potential of a VRT technology with respect to neuroplasticity and motor learning, it is not enough to characterize children’s enjoyment of the activity. While motivation is an extremely important factor in neuroplasticity, a child may enjoy an activity that does not provide the requisite feedback or complexity for instance, to be therapeutically relevant or functionally beneficial. As such, this review suggests the need to evaluate VRTs with respect to all four elements considered important for neuroplasticity and motor learning to fully assess the potential value of these technologies and the mechanisms by which they may elicit functional improvements. To accomplish this, new tools are needed particularly for the measurement of feedback and for characterizing task complexity. While the latter, for instance, is often measured with respect to fatigue and level of difficulty, even when these measures are completed, it is unclear how they should be interpreted and what the optimal target values should be. New measures are needed for evaluating cognitive complexity of tasks. As well, we feel it would be prudent to also measure variability of task performance across randomization of tasks during sessions.

ii. **Create more standard evaluation tools.** 81% of studies used customized tools to evaluate VRT technologies. While customized tools enable in-depth evaluation of issues that may be unique to the specific technology, it was observed that in many cases, the same issues were likely of interest across studies. This suggests that there may be value and a need for the design of more standard questionnaires that address the key factors of neuroplasticity and also, elements of system usability and user interactions. This would better enable comparisons between studies which could then lead to the design of improved systems and platforms. The availability of standard tools could also reduce the resources, time, and expertise required to evaluate VRT technologies prior to clinical deployment. It is suggested that the development of these standardized tools should be a collaborative effort to ensure their general applicability. As a first
step, it may be valuable to explore potential tools that have been used in other research areas; for example, the System Usability Scale (SUS) [36], the Intrinsic Motivation Inventory [37], the Physical Activity Enjoyment Scale (PACES) [38], the OMNI Scale of Perceived Exertion [38], or the Borg Scale of Perceived of Exertion [39], for instance.

iii. **Use a balanced mixture of objective and subjective measurement tools.** The majority of measures identified in this review used subjective as opposed to objective tools. Subjective measures are essential for capturing users’ perceptions and interpretations. Objective measures can be used to corroborate subjective measures and can also reduce the number of questionnaires that must be completed. Objective measures may be particularly relevant for studies involving young children who may not fully understand subjective questionnaires. Future VRT technologies should be constructed with these evaluations in mind as the capacity for seamlessly collecting data on usage and skills acquisition for instance is a valuable advantage of these systems. One potential reason game statistics are not reported more often in VRT studies is likely the common usage of commercial game systems (e.g. Nintendo® Wii) that are not intended for rehabilitation purposes and as such, do not record information of value to their evaluation. Nevertheless, greater utilization of the game environment to study and manage task complexity and practice/repetition in particular, is desirable.

iv. **Conduct multi-perspective evaluations.** There are many key stakeholders in the design and evaluation of VRT technologies that can include patients, therapists, parents, and even siblings/friends, if social play is being encouraged. Understanding interactions of each of these user groups with the VRT technology as well as interactions between key stakeholders (e.g. therapist-child dynamics) in the use of these technologies is important to the overall evaluation of these systems.

v. **Conduct context specific evaluations.** VRT technologies may be intended for use in the clinic, school, home, and/or in multiple contexts. Ensuring that a standardized toolkit has options for measurement that are appropriate for the context in which the
VRT is intended to be used is an important consideration of note. The importance of evaluating VRT technologies within their intended context should be emphasized.

2.4.3 Recommendations towards a Standardized Framework

How are VRT Technologies being evaluated?

The emergence and study of different VRT technologies is increasing rapidly, yet, clinical implications are still inconclusive [4]. The majority of studies included in this literature review evaluated the ability of VRT technologies to elicit changes in children’s functional outcomes as modeled in Figure 4. The benefits of the current paradigm are that evaluations are based on the overarching goal of therapy to achieve functional improvements and that it becomes possible to understand the translation of skills learned in VRT into activities of daily life. However, the issues with this methodology are that it is not conducive to the iterative design of technologies, especially important in the early stages of technology development; also, it is difficult to understand the “active ingredients” of an intervention. “Active ingredients” are defined by Levac et al. [9] as “reasons why a treatment is expected to be effective.” This review specifically posited that research in the VRT field has revolved around determining if functional outcomes are realized, without really delving into the question of why they work (or don’t work).

![Figure 4: Current paradigm for evaluating VRT technologies](image-url)
How should VRT Technologies be evaluated?

It is increasingly important to evaluate VRT technologies more comprehensively and at an earlier stage of their design process in order to understand the “active ingredients” that they may be realistically expected to offer. This understanding is key to guiding implementation and maximizing potential success when clinically applied. Understanding these “active ingredients”, particularly as relates to neuroplasticity, will help to demystify the “black box” by which therapies change functional outcomes. To address these issues, we propose a 3-staged evaluation framework for facilitating the iterative design process and maximizing the success of VRT interventions as shown in Figure 5.

1. System Design and Usability Testing. In this first phase of evaluation, we propose that interactions between the child, therapist, and technology should be investigated with respect to system usability and performance. In this first phase, system validation studies should be carried out to ensure that all software, hardware, etc. performs well in the specific context and target population for which it is intended. Usability testing should be performed both with children and with therapists. If there are additional users expected to interact with the system, usability testing should also include these stakeholders. Once the system is adequately accurate and usable, evaluations may proceed to the second phase.

2. Feasibility Testing to Determine “Active Ingredients”. In the second phase, a feasibility study should be performed to specifically evaluate the VRT technology for its ability to elicit the “active ingredients”. In the case of VRT technologies designed to improve motor learning and neuroplasticity, the 4 factors of interest are: motivation, feedback, task complexity, and practice. Measures and evaluation tools may be chosen from a number of tools identified in the previous literature review and should evaluate for all 4 factors. Based on these evaluations, two outcomes can be expected:

a. If the system does not successfully provide the “active ingredients”, a design iteration should occur. Areas for improvement identified from the feasibility study should be used to redesign and improve the system. Phase 1 of the VRT
evaluation framework should be repeated following the redesign prior to repeating Phase 2.

b. If the system is able to significantly promote the factors associated with neuroplasticity in children during therapy, iteration is not necessary, and evaluation should proceed to phase 3.

3. Evaluating Clinical Impact. In the third, and final, phase, clinical impact of the system can be evaluated to understand the therapeutic effects of the system on functional outcomes in children. Functional outcomes are well-documented and may be chosen from one of many available systematic reviews. [4] [5] [6] [7] [8] [22] [40] [41] [42]

**Proposed framework for VRT evaluation**

![Proposed framework for VRT evaluation](image)

**Figure 5: Proposed iterative 3-stage VRT evaluation framework**

This framework fits within the typical engineering design cycle as shown in Figure 6. The first step, understanding, includes preliminary design activities such as defining a problem and identifying user needs through methods such as focus groups, shadowing, task analysis,
etc. Once a problem has been identified, a potential solution is generated in the design step. The framework developed here would fit into the evaluation step of the overall design cycle.

Figure 6: Iterative design cycle adapted from [43]. Dotted line indicates the step in the cycle where the developed framework should be applied.

This framework allows for evaluation of both the technology and the intervention. Phase 1 and 2 of the framework allows for testing of the technology which may or may not occur within the therapeutic context of the intervention. However, the 3rd phase allows for evaluation of the overall VRT intervention.

Phase 1 consists of usability testing and system characteristics evaluation. Usability can be assessed through short-term study, through single sessions with a small, representative sample of users. System characteristics may require a greater number of participants and testing sessions to adequately account for movement variability in the design. Phase 2 consists of feasibility testing. Ideally, feasibility testing should take place over many sessions to mitigate “transient” responses to the technology. For example, if participants are only presented with the VRT technology in one session, they may highly enjoy using it; however,
over time, the novelty of the technology may wear off and interest may decrease. While resources may limit the ability to conduct longitudinal feasibility studies prior to clinical evaluation, the success of the eventual intervention will be more likely if temporal considerations are at this second stage of the framework.

Potential significance of the proposed framework

This proposed framework follows the iterative and user-centred design (UCD) process widely-accepted in the technology industry. This framework is novel in its specific application to the design and evaluation of VRT technologies and integration of key factors in motor learning and neuroplasticity.

Unlike current methods of evaluating VRTs, the proposed framework is more conducive to early testing and iterative design of new interventions. Additionally, it provides a means of performing root cause analysis for the successes/failures of technologies where promotion of motor learning has not been observed in functional outcome studies. In some cases, functional improvements may be realized but not through mechanisms of motor learning/neuroplasticity (e.g. through endurance or strength building). This is also important to identify and understand.

Another benefit of the proposed framework is that it allows specificity in terms of areas of improvement of the technology (e.g. it may identify that the feedback provided by the system was not clearly understood by the user for instance). Additionally, initial phases of evaluation are not resource intensive. Note: usability testing can successfully identify the majority of usability issues with as few as 4 to 5 representatives of a specific user group. This means that potential problems can be identified and rectified rapidly and early on prior to the more large-scale, resource intensive clinical evaluations associated with Phase 3.

Who should use the framework?

This framework is applicable for researchers, game designers, health professionals, and engineers working in the area of VRT technology development and evaluation. Collaborative design and evaluation efforts are recommended between the aforementioned stakeholders. However, the onus for championing the implementation of the evaluation framework would
likely be on the researcher interested in investigating the potential of a new VRT and the clinicians who will be applying it.

2.4.4 Limitations and Merits

Limitations

While structured and methodical in nature, this system was not an exhaustive, systematic review and therefore, may not have identified all relevant articles. The tools provided are based off of the tools previously used in the current literature and do not include other measures that may exist, but have not yet been utilized in the context of motor rehabilitation in children with disabilities. Also, the current literature base does not allow for strong recommendations with regards to validated or standard measures, nor were evaluation measures consistently well-described and documented. Unfortunately, due to the heterogeneity of the studies identified in the review, it is difficult to provide specific recommendations for promising tools. Future work is needed to further develop standard measures and approaches for evaluation in this area that enable comparisons between studies.

While the framework developed here is conceptually significant, it is still preliminary and we feel that it should be further developed prior to widespread adoption in the research community. For example, use of the framework is heavily reliant on a standardized toolkit, which will need to be developed. Additionally, the framework may not work for all types of technologies being introduced in therapy. Researchers and clinicians may find there are two specific types of technologies that may be integrated: customized and off-the-shelf. Customized technologies may be developed in-house according to specific therapeutic needs of an individual or population, whereas off-the-shelf technologies (e.g. Nintendo® Wii) are technologies that are pre-existing and have not specifically been designed for therapeutic use. The use of the framework for each of these technology contexts may be slightly different. While customized technologies may be iteratively redesigned, off-the-shelf technologies are less flexible and, therefore, it may be difficult to iterate their redesign. Future development of the framework should address both types of technologies and provide recommendations for each of these contexts.
Merits

The tools and evaluation framework provided herein are timely, considering the increase in studies involving VRT, ICP, ART, etc. in the past few years. It provides important directions for future work in the fields of motor rehabilitation in children with disabilities. The tools are also a valuable planning resource for designers, researchers, and clinicians embarking on the development and/or evaluation of technologies in these fields. Finally, it provides a methodological framework for better ensuring that these technologies are aligned with the fundamental framework of neuroplasticity for promoting aspects of motor learning. We expect that the use of the proposed framework and toolkit will ultimately allow for more comprehensive evaluations.

2.5 Conclusion

The study of therapy-augmenting technologies such as VRT is increasingly rapidly, yet, the findings of this review show that there are few studies that have adopted a holistic approach to evaluating low-level aspects of the interaction of children and therapists with these technologies with respect to factors affecting motor learning and neuroplasticity. This paper proposed a framework for how VRT technologies should be evaluated that adopts a phased approach conducive to the user-centered design cycle. It provides recommendations for the development of a toolkit of measures that is more comprehensive, relies more on standard measures, balances subjective and objective perspectives, is context-specific, and solicits perspectives from multiple stakeholders. It is hoped that the recommendations presented in this paper may lead to a better understanding of the “active ingredients” of motor learning therapies based on VRT technologies.
Chapter 3

3 Using the Musical Steps augment-reality system toward heel contact in toe-walking children: A case study

3.1 Introduction

3.1.1 Motivation

As outlined in the previous chapter, Virtual Reality Therapy (VRT) technologies are increasingly being developed and evaluated for their therapeutic effects in children with motor disabilities. However, there is insufficient evidence to suggest that these technologies have been successful in eliciting motor learning as evidenced by functional outcomes. [4] The issues with the current paradigm for VRT technology evaluation have been identified in the previous chapter and a novel framework was developed to aid in understanding the “active ingredients” leading to functional changes in children undergoing therapy. The development of this framework led to the following research question: How can this framework be applied to a new VRT intervention? To answer this question, a case study was performed to investigate the use of the framework in the first iteration of the development and evaluation of a novel “Musical Steps” VRT technology.

3.1.2 Therapeutic Goal

The ability to walk in a safe and efficient manner is important for children to participate in activities of daily living and social life. However, children with functional impairments often realize restrictions in these aspects of their lives. [49] One specific pathological gait pattern observed in children with functional impairments is equinus gait, or “toe-walking”. Toe-walking refers to pathological gait patterns where typical heel contact is not achieved during the stance phase of steps and toe-toe patterns are observed. [50] Toe-walking is very common in all children under the age of 24 months. [51] However, this phenomenon is only considered pathological in children past the age of four, once gait maturation has occurred. [52] Toe-walking is one of the most prevalent gait problems in children with cerebral palsy (CP) [53];
it observed in 61% of children with CP [54], which affects 2-2.5 out of every 1000 children [55]. Within this group, equinus gait may either be considered fixed or dynamic. Children with fixed equinus gait have static contractures of the triceps surae muscle. This restricts them from achieving the necessary range of motion in dorsiflexion for achieving heel contact when walking. However, children with dynamic equinus are capable of achieving adequate dorsiflexion, with no static contractures, yet exhibit contractures of the gastrocnemius and soleus muscles during walking; thus, they exhibit toe-walking. [56] Toe-walking may also be considered “idiopathic” in children when it is observed in children with no neurological problems. [57] Lastly, toe-walking has been exhibited in 19% of children with autism spectrum disorder (ASD), which may be more related to behavioural preference rather than physical disability. [54]

The implications of toe-walking in children involve the reduction of safe and efficient ambulation, affecting both progression and stability during walking. The consequences of not using a typical gait pattern involve secondary postural problems, other gait deviations, reduced physical activity, and difficulty in performing advanced motor skills and high level sports activities. [57] Children may experience a diminished endurance and distance travelled. [58] In terms of stability, the reduced likelihood of heel contact for each step limits the base of support when balancing on the foot during the stance phase. As a result, children lose the ability to smoothly translate their bodies over their standing foot, and may also experience inadequate foot clearance during swing phase [59], potentially leading to trips and falls.

3.1.3 Physical Therapy for Toe-Walking Children

In treating toe-walking in children, physical therapy is often used alongside various other treatment methods such as botulinum toxin injections, serial casting, and orthoses [58], with an overall goal to facilitate safe and efficient mobility in children exhibiting toe-walking. Physical therapy has been used in facilitating the prevention and correction of equinus deformity, promoting a base of support, training of skills, and improving the efficiency of the patient’s gait. [58] While physical therapy may involve many other exercises, including stretching and strengthening, gait training is one important component of a child’s physical therapy. However, gait training can be challenging for children due the nature of their
impairments. Since achieving heel contact, or walking in general, may be stressful and frustrating for these children, they may experience decreased motivation and enjoyment during therapy sessions. This may slow their overall learning and progress in therapy, since motivation is an important factor affecting a child’s motor learning.

Currently, therapists utilize three main motor learning strategies in their therapy sessions: verbal instructions, practice, and verbal feedback. Verbal instructions may be used to promote the therapeutic goal by coaching children on specific aspects of the task at hand, while verbal feedback of a child’s performance may be qualitative (positive or negative) or quantitative. [12]

3.1.4 Auditory Biofeedback

In line with current practice of therapists, operant conditioning may be used for training motor skills in children with CP. [60] [61] Skinner [62] described operant conditioning as the behavioural learning through reinforcement of consequences of behaviour. Positive reinforcement is the use of a rewarding stimulus delivered after a desired response is observed in the subject to change their behaviour such that the desired response will occur with a greater frequency. [62]

In physical therapy, biofeedback can relay information to the patient regarding their movements, which can potentially help in the control of these movements. Previous work has investigated the use of auditory biofeedback in gait training for various patient populations including cerebral palsy. [63] [64] [65] [66] Two of these studies [63] [65] investigated the effects of auditory feedback on improving ankle dorsiflexion in children with promising results. In these studies, a simple switch on the heel triggered a buzzer when heel contact was made. More broadly, auditory stimulation has also been shown to influence the time-keeping of rhythmic patterned movements. In particular, Rhythmic Auditory Stimulation (RAS) can improve gait characteristics, such as velocity and stride length in children with CP. [21] In general, physical therapies which integrate music/sound as a motivator have resulted in improved outcomes in motor function and enjoyment. The use of music as a part of PT technique has been referred to as ‘musical therapy’.
3.1.5 Gait Analysis Technologies

Many technologies have been proposed for objective measurement of gait metrics for children participating in gait training therapies. The majority of these technologies however, are not practical for use in everyday clinical settings by physical therapists (PTs) or physical therapy assistants (PTAs). Factors facilitating implementation of gait technologies in clinical settings include usability, portability, cost-effectiveness, ability to detect patient-specific gait events in real-time, detect spatial displacement, and multi-purpose use across therapies (ex. walking, stair climbing, standing). The biomechanical standard for gait analysis (i.e. infrared camera systems such as the Vicon®) is expensive, stationary, and time consuming to set up. Other commercial systems have been proposed that utilize pressure sensitive mats (i.e. GaitRite®), but these often do not meet the needs of therapists with regards to portability and user-friendliness. Additionally, when therapists must walk alongside children to provide support during the therapy sessions, these technologies are unable to distinguish between the therapist and child in real-time. Thus, gait analysis technologies are not commonly used in physical therapy sessions. Alternative approaches that have been proposed for gait measurement include accelerometers [67], pressure sensors [68], capacitive sensors [69], resistive sensors [70], Radio Frequency Identification [71], electromagnets [72], and computer vision [73].

One of the simplest technologies for providing auditory feedback in response to heel contact is force sensing resistors (FSRs) attached to the sole of the shoe. FSRs are portable and provide a reliable “on/off” switch as needed to detect heel contact [74]. Simple “on/off” foot switches have been used in the past to reward heel dorsiflexion [63] [65]. However, these studies did not incorporate any algorithmic logic to ensure that the auditory feedback was actually rewarding heel contact associated with progressive steps forward. Inertial sensors are another promising technology given their ability to provide a wealth of information regarding acceleration and velocity profiles during gait. These technologies, while widely used for describing typical gait patterns [75] [76] [77] [78] [79] [80] [81], need further study and development for the analysis of pathological gait in children, particularly in the area of toe-walking.
3.1.6 Study Objectives

The main goal of this study was to demonstrate how the novel VRT evaluation framework (Figure 5) can be applied to the evaluation of a new VRT technology. To accomplish this, we:

(i) designed ‘Musical Steps’, a VRT technology to address “toe-walking” in gait therapy, and
(ii) evaluated ‘Musical Steps’, applying the developed framework. The Musical Steps system was aimed at augmenting the physical therapy environment through the integration of musical feedback as a means of positive reinforcement. The specific therapeutic goal of interest was increased heel contact during ambulation and the target population was children exhibiting toe-walking. Additionally, the system provided a portable means of measuring and tracking objective measures of gait function (e.g. number of steps) for clinicians. It was hypothesized that the system would help motivate and reward children’s engagement in therapy. In evaluating this system, we sought to understand its usability and validity, assess its potential for positively affecting factors of neuroplasticity, and identify areas for improvement or redesign.

![Proposed framework for VRT evaluation](image)

**Figure 7: Previously developed iterative 3-stage VRT evaluation framework**
3.2 Methods

3.2.1 Musical Steps – System Design

Sensors.

While not validated for use with toe-walking children in gait therapy sessions, Smith et al. [74] demonstrated that 94.5% of steps can accurately be detected using only FSRs. As such, it is hypothesized that FSRs may be a reliable approach to detect heel contact during gait therapies. Two FSR® 406 sensors (Interlink Electronics Inc., Camarillo, CA, U.S.A.) were attached to the sole of each shoe at the heel and toe. Two Shimmer Wireless Sensor Units, each equipped with a 9DoF Daughter Board and Expansion Board, were used to collect and wirelessly transmit the data (Realtime Technologies Ltd., Dublin, Ireland). The FSRs were attached to the analog inputs of the Expansion Board for both Shimmer units as shown in Figure 9. The connection wires were equipped with reusable connectors which allowed FSRs to easily be replaced for every session so that any wear (ex. tearing, scratches, etc.) did not interfere with the system accuracy or integrity. The Shimmer units were also set to collect gait kinematic data through the 3-axis accelerometers, 3-axis gyroscopes, and 3-axis magnetometers. The sampling rate was 100 Hz. These data were transmitted, in real-time, wirelessly over Bluetooth® to a laptop which recorded data via a custom software application in Microsoft Visual C++ 2012 (Microsoft Corp., Redmond, WA) that had been previously developed [37] and was modified to collect FSR data for the purposes of this study. Shimmer sensor calibration was performed using the Shimmer 9DOF Calibration v2.3 (Realtime Technologies Ltd., Dublin, Ireland) software. The Shimmer units were attached such that the x-axis was collinear with the coronal axis, the y-axis with the sagittal axis, and the z-axis with the transverse axis as demonstrated in Figure 8 below.
A rule-based algorithm was designed to identify and reward heel contact associated with progressive steps. This algorithm was based on a previous design by Popovic et al. [82] (see Figure 10 below). The algorithm was designed specifically for the case of straight, over ground walking and rewarded heel contact made in the course of a step. Of note, only one heel strike
was rewarded per step. The timing of the heel contact was not critical. While, ideally, steps should be initiated with heel strike, this is considered an advanced skill. For gait training with young children, the preliminary therapeutic goal is to improve heel contact in general. The algorithm was designed a priori based on typical gait patterns. Validation with respect to its accuracy in the target population is paramount given that the range of pathological gait observed in children may introduce greater variability than it was designed for.

![Algorithm A logic flow diagram using only FSR data to detect heel contact events](image)

**Figure 10:** Algorithm A logic flow diagram using only FSR data to detect heel contact events

**Musical Feedback.**

To the best of the author’s knowledge, there has been no indication in the literature as to which types of feedback (ex. Metronome clicking, buzzing sounds) are the most rewarding for auditory biofeedback in gait. It was hypothesized that children would enjoy musical sounds
more than the other types of auditory feedback. Therefore, the system was designed such that when the algorithm detected heel contact, the laptop would output notes sequentially associated with a simple children’s song, “Twinkle, Twinkle, Little Star”. This song was chosen given its familiarity and because of its even 4/4 rhythmic meter, which is not expected to influence the gait symmetry of children.

Clinician Interface.

A prototype interface was designed to present gait metrics to therapists such as the number of steps taken, percentage of steps with heel strikes, total number of heel strikes, average step time, and cadence. The prototype interface did not link with the sensors in real-time, but it did have the look and feel of the final interface, and allowed for full interactivity. Three key functionalities were identified: (1) start a session, (2) end a session, and (3) review session results. To support this workflow, three main screens and one confirmation modal window were designed using standard design heuristics.

Figure 11: Screens and workflow of the developed prototype clinician interface
3.2.2 Evaluation: Application of Novel VRT Framework

Ethics approval for the studies outlined below was obtained from the research ethics board at Holland Bloorview Kids Rehabilitation Hospital and the University of Toronto.

Study Design

An AB single-subject research design was utilized with randomization of the order between the baseline, phase A (i.e. gait training with no feedback), and the intervention, phase B (i.e. gait training with musical feedback). Participants’ study sessions were randomly ordered such that 2 participants received musical feedback in their first session, while 2 received musical feedback in their second session. The study sessions occurred no more than 2 weeks apart. Follow-up sessions were scheduled with individual therapists to review the user interface with respect to its usability.

Participants

Children enrolled in gait therapy for toe-walking were recruited from the Child Development Program at Holland Bloorview Kids Rehabilitation Hospital and the Bloorview School Authority. Children were included only if they met the following criteria: (i) enrolled in physical therapy that includes a gait training component at Holland Bloorview Kids Rehabilitation Hospital, (ii) between the ages of 4 and 8 years, (iii) capable of understanding simple two-step gross motor instructions, (iv) capable of answering questions in English, (v) met the walking functional ability as identified by the physical therapists (ex. Gross Motor Function Classification System (GMFCS) Levels I to III for children with CP), (vi) ambulatory and capable of walking for 5 minutes on a horizontal surface with or without the use of assistive devices, and (vi) exhibit toe-walking and have a physical therapy goal to increase heel contact. Informed consent and assent were obtained from the children’s parents and from each child, respectively.

Convenience sampling was used to recruit therapists for the study who met the following criteria: (i) must be a physical therapist or physical therapy assistant assigned to the child participant, (ii) must work with the child population of interest, (iii) must have experience with
gait therapy, and (iv) must be working at Holland Bloorview Kids Rehabilitation Hospital. Informed consent was also obtained from each therapist that participated.

Lastly, convenience sampling was used to recruit a therapist to perform a blinded review of video recordings of the children’s interactions with the Musical Steps system. Inclusion criteria included: (i) must be working at Holland Bloorview Kids Rehabilitation Hospital, (ii) must have experience with gait therapy, and (iii) must be either a physical therapist or physical therapy assistant. Therapists were excluded if they were involved in any of the study’s therapy sessions. Informed consent was obtained from the therapist.

Protocols

Child Interaction with Technology

Each child participated in 2 study sessions. The protocol was identical in both of the sessions. Informed consent was obtained at least 1 week prior to study sessions from the children’s parents. At the start of their first session, informed assent was obtained from each child. Demographic information was also obtained for both the child (APPENDIX A) and the therapist (APPENDIX B). The sensors were then attached to each of the child’s shoes. For each shoe, the Shimmer units were attached to the posterior aspect (heel) as shown in Figure 12 and the FSRs were placed under the heel and toe as shown in Figure 13. Participants’ lower limbs and feet were video recorded using a Microsoft LifeCam NX-6000 Webcam (Microsoft Corp., Redmond, WA) at 30 fps using the native Microsoft LifeCam software. The same laptop was used to simultaneously collect the video recording. Participants’ facial expressions and reactions were also recorded using a Sony DCR-SR85 (Sony Corp., Tokyo, Japan) handheld digital camcorder at 30 fps.
Once all sensors were attached, video recording and sensor sampling was initiated. The webcam was used to record a synchronization notification which appeared on the laptop screen. This provided a marker in the video data to synchronize with the sensor recordings. Testing took place in a hallway in Holland Bloorview Kids Rehabilitation Hospital where start and endpoints for a 10-m runway were marked on the ground (Figure 14).
Figure 14: Testing environment for the gait sessions was a hallway with a 10-meter runway marked on the floor

The therapists were provided with a standard list of tasks for the children to perform and began the gait therapy session. Children were given 5 minutes to complete the sequence of tasks which included: i) walk 30 m forward at a comfortable pace, ii) walk 10 m backward at a slow pace, iii) walk 10 m forward at a fast pace, iv) walk 10 m forward at a slow pace, and v) walk 10 m backward at a comfortable pace. Children were then allowed to take a break. Once the child was ready to start again, the therapist guided them through the same sequence of tasks again. During both of the sessions, PTs and PTAs provided their standard level of care with respect to physical assistance, if necessary, and praise comments not specific to encouraging heel contact, such as "great work", "keep it up", or "I like the way you are working with the
feedback”. Upon completion of the two circuits, the sensors were removed from the child’s shoes. After each session, children were asked to complete the appropriate questionnaires as detailed in the Evaluation Measures and Tools section to follow. Questionnaires were completed with the help of the researcher and PT or PTA, if necessary.

A blinded observer (an independent PTA) reviewed muted video recordings of the AB study sessions with each of the 4 children and to complete a questionnaire commenting on the child’s behaviours during each session. The blinded observer was given access to all the session recordings for analysis over a span of two non-consecutive days. Once all sessions had been analyzed, the PTA was informally interviewed to gather further feedback on the Musical Steps study/sessions.

Therapist Interaction with Technology

A laptop running Windows XP (Microsoft Corp., Redmond, WA) was utilized in the interface usability testing with therapists. A Sony DCR-SR85 (Sony Corp., Tokyo, Japan) handheld digital camcorder was used to record video of the laptop’s screen and audio data.

No training was provided prior to beginning the session. Each therapist was told that the prototype was a non-functional user interface in terms of linking to sensors and that the data displayed was fictitious. However, they were also informed that in the future, the system would be able to immediately display gait metrics during therapy sessions for the therapist and child to review. Each therapist was told that the purpose of the testing was to gauge their interest and opinions about the usability and functionality of the prototype.

Therapists were asked to complete 3 tasks (start a session, save a session, review a session). There was only one way to complete each task. Therapists were asked to complete the tasks while “thinking aloud” [85] to provide insights into their thought process, how easy the interface was to navigate, like/dislikes about the interface and any missing information or functionality. During the tasks, the researcher also asked some specific questions about their understanding of what the current system state was (e.g. are the sensors on?) and their opinions of information layout and functionality. At the end of the sessions, therapists were debriefed and an unstructured interview was performed to understand their general thoughts on the
Musical Steps system as a whole, including the prototype interface, the sensors, system setup time, musical feedback, practicality for clinical use, preference for type of display screen (e.g. smartphone), etc. Once the interview was complete, the therapist completed a post-test questionnaire to summarize their views of the Musical Steps system and also provide any additional feedback.

**Evaluation Measures and Tools.**

**Phase 1: System Performance and Usability Testing**

**System Validation**

The system performance was evaluated via the accuracy and percentage of false positives of the heel step detection algorithm summarized in Table 7.

**Child and Therapist Usability Measures**

User interactions with the system were evaluated via measures of user preference, usability, and general feedback are summarized in Table 8.

**Phase 2: Feasibility Studies**

**Motivation Measures**

Motivation was evaluated using the following measures: enjoyment/fun, distraction, and engagement. The summary of tools used to evaluate each of these measures is provided in Table 9.

**Feedback Measures**

Feedback was evaluated by measuring interpretation of the system feedback. The summary of tools used to evaluate the feedback usability is provided in Table 10.

**Practice Measures**

Practice was evaluated through measuring the achievement of the therapeutic goal and the tools for evaluation are provided in Table 11.
Task Complexity/Difficulty

Task complexity was evaluated using the following measures: exertion and fatigue. The summary of tools used to evaluate each of these measures is provided in Table 12.
Table 7: Summary of measures and tools used to evaluate the system design and usability testing

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Usability testing with children</th>
<th>Usability testing with therapists</th>
<th>Sessions Reported</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Heel Step Detection Algorithm</strong></td>
<td>Gait metrics</td>
<td>☑</td>
<td></td>
<td>All</td>
<td>The average accuracy for detecting heel Steps for overground walking was determined using data across all participant gait sessions.</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Heel Step Detection Algorithm</strong></td>
<td>Gait metrics</td>
<td>☑</td>
<td></td>
<td>All</td>
<td>The average % of false positives detected for overground walking was determined using data across all participant gait sessions.</td>
</tr>
<tr>
<td><strong>Percentage of False Positives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Summary of measures and tools used to evaluate user interaction with the Musical Steps system

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Usability testing with children</th>
<th>Usability testing with therapists</th>
<th>Sessions Reported</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>User Preference</strong></td>
<td>Child Post-Test Questionnaire</td>
<td>☑</td>
<td></td>
<td></td>
<td>Using a dichotomous Yes/No response, children were asked to rate their agreement with the following statements:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• “I liked the sounds.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• “The sounds were too loud.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• “The sensors were annoying to wear.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Children were asked to answer the following open-ended questions:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• “Are there any other sounds you would like to hear?”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• “What did you like best about Musical Steps?”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• “What didn’t you like about Musical Steps?”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• “How could we make Musical Steps better?”</td>
</tr>
</tbody>
</table>
## Usability

<table>
<thead>
<tr>
<th>Task</th>
<th>Completion Rate</th>
<th>Usability Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Therapist Post-test Questionnaire</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Tasks completed by therapists during usability testing of the prototype interface were marked as completed or failed. Rate of completion was recorded for the following tasks: (i) Start a session, (ii) End a session, (iii) Navigate to session summary page.

On a 5-pt Likert scale (1=Agree, 5=Disagree), therapists reported their agreement with the following statements:

- “The performance data (e.g. number of steps taken, cadence) from the Musical Steps system could be helpful to me.”
- “I would use this system in regular therapy sessions if it was available.”
- “I think the Musical Steps system is a useful tool to use in therapy sessions.”
- “The benefit of the system does not justify the added setup time in using it.”
- “The system is too difficult to use.”

Therapists were asked to provide feedback on the statement, “I would improve the system by…”

The researcher recorded all usability issues experienced during the setup and usage of the Musical Steps system.

## General Feedback

<table>
<thead>
<tr>
<th>Feedback</th>
<th>Unstructured Interview</th>
<th>Usability Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>After Blinded Video Analysis Completed</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Therapists were probed for their general thoughts and feelings toward the Musical Steps system both during and after usability testing. Topics of potential discussion included the instrumentation required, setup, interface preferences, usefulness of the auditory feedback, preferences for future directions toward clinical implementation.

An unstructured interview with the therapist was conducted to gather their summative opinions of the study after completion of analysis of all videos for all study gait sessions.

Therapist was asked to provide additional comments for all session videos.
Table 9: Summary of measures and tools used to evaluate motivation

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Response Type</th>
<th>Usability testing with children</th>
<th>Sessions Reported</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Enjoyment/ Fun</strong></td>
<td>Smileyometer [86]</td>
<td>5-pt pictorial scale (1=Awful, 5=Brilliant)</td>
<td>✓</td>
<td>All</td>
<td>Children were asked, “How fun was it to do this session?”</td>
</tr>
<tr>
<td></td>
<td>Therapist Post-test Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>After musical feedback</td>
<td>Therapists were asked to rate their agreement with the statement, “The client enjoyed the Musical Steps system.”</td>
</tr>
<tr>
<td></td>
<td>Blinded Video Analysis Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>All</td>
<td>Therapist was asked to rate their agreement with the statement, “The child appeared to enjoy the therapy session.”</td>
</tr>
<tr>
<td></td>
<td>Child Post-test Questionnaire</td>
<td>Dichotomous Yes/No response</td>
<td>✓</td>
<td>After musical feedback</td>
<td>Children were asked to rate the agreement with the statement, “I liked walking with the sounds more than without.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Children were asked, “Do you want to make the sounds happen again?”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Children were asked to rate the agreement with the statement, “I had fun walking”.</td>
</tr>
<tr>
<td><strong>Distraction</strong></td>
<td>Therapist Post-test Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>After musical feedback</td>
<td>Therapists were asked to rate their agreement with the statement, “The sounds distracted the client and interfered with the therapy sessions in a negative way.”</td>
</tr>
<tr>
<td></td>
<td>Blinded Video Analysis Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>All</td>
<td>Therapist was asked to rate their agreement with the statement, “The child seemed distracted during the therapy session.”</td>
</tr>
<tr>
<td><strong>Engagement</strong></td>
<td>Blinded Video Analysis Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>All</td>
<td>Therapist was asked to rate their agreement with the statement, “The child appeared engaged/motivated during the session.”</td>
</tr>
</tbody>
</table>
Table 10: Summary of measures and tools used to evaluate feedback

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Response Type</th>
<th>Usability testing with children</th>
<th>Usability testing with therapists</th>
<th>Sessions Reported</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sense of Agency</strong></td>
<td>Therapist Post-test Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td></td>
<td></td>
<td>Therapists were asked to rate their agreement with the statements, “The client understood how the Musical Steps system worked and how to activate the sounds” and “The sounds helped the client focus more on achieving heel strikes.”</td>
</tr>
<tr>
<td></td>
<td>Child Post-test Questionnaire</td>
<td>Open-ended</td>
<td>✓</td>
<td></td>
<td></td>
<td>Children were asked, “How did you make the sounds play?”</td>
</tr>
<tr>
<td></td>
<td>Dichotomous Yes/No response</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>Children were asked to rate the agreement with the statement, “The musical walking made me want to use my heels when I walked.”</td>
</tr>
</tbody>
</table>

Table 11: Summary of measures and tools used to evaluate practice

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Response Type</th>
<th>Usability testing with children</th>
<th>Sessions Reported</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Achievement of Therapeutic Goal</strong></td>
<td>Gait metrics</td>
<td>Sensor + Video data</td>
<td>✓</td>
<td>All</td>
<td>Calculation of the Percentage of total steps taken with heel contact</td>
</tr>
<tr>
<td></td>
<td>Therapist Post-test Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>After musical feedback</td>
<td>Therapists were asked to rate their agreement with the statement, “The client performed better (e.g. achieved more heel strikes) with the musical feedback.”</td>
</tr>
<tr>
<td></td>
<td>Blinded Video Analysis Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>All</td>
<td>Therapist was asked to rate their agreement with the statement, “There was evidence that the use of heel strikes was encouraged.”</td>
</tr>
</tbody>
</table>
## Table 12: Summary of measures and tools used to evaluate difficulty

<table>
<thead>
<tr>
<th>Measure</th>
<th>Tool</th>
<th>Response Type</th>
<th>Usability testing with children</th>
<th>Sessions Reported</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exertion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P-CERT [87]</td>
<td>10-pt pictorial scale</td>
<td>✓</td>
<td>All</td>
<td>Children were asked, “How difficult was it to do this session?”</td>
</tr>
<tr>
<td></td>
<td>Therapist Post-test Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>After musical feedback</td>
<td>Therapists were asked to rate their agreement with the statement, “The client exerted themselves more when there was musical feedback.”</td>
</tr>
<tr>
<td></td>
<td>Blinded Video Analysis Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>All</td>
<td>Therapist was asked to rate their agreement with the statement, “The child appeared to exert themselves to the best of their ability.”</td>
</tr>
<tr>
<td><strong>Fatigue</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blinded Video Analysis Questionnaire</td>
<td>5-pt Likert scale (1=Agree, 5=Disagree)</td>
<td>✓</td>
<td>All</td>
<td>Therapist was asked to rate their agreement with the statement, “The child was fatigued by the therapy session.”</td>
</tr>
<tr>
<td></td>
<td>Child Post-test Questionnaire</td>
<td>Dichotomous Yes/No response</td>
<td>✓</td>
<td>After musical feedback</td>
<td>Children were asked to rate the agreement with the statement, “I felt tired.”</td>
</tr>
</tbody>
</table>
3.2.3 Data Analysis

The number of steps and the number of heel steps were determined via a combination of video and FSR data. Videos were reviewed to obtain the timestamps of all actual steps taken by the participants. Since the algorithm was designed specifically for the case of straight, over-ground walking, steps were only included if the participant was clearly making a full transfer of weight in forward or backward progression between feet. Time frames in the video were excluded if (i) the participant was not taking progressive steps (ex. standing), (ii) the participant was out of view of the camera, (iii) a sensor came off of the child’s foot, or (iv) the participant was taking a break between the two walking circuits.

The timestamps of all steps were compared against FSR data to obtain the number of steps where heel contact occurred. The percentage of heel steps was determined using the ratio between the number of heel steps and the total number of steps. A quasi-confusion matrix (Table 13) was developed to aid in validation of the heel step detection algorithm. This matrix is considered a quasi-confusion matrix because true negatives (i.e. heel steps that did not occur and were not detected) were not included in the model, as this condition is infinitely true and does not fit within the context of this evaluation.

**Table 13: Quasi-Confusion Matrix for Heel Step Detection Algorithm Evaluation**

<table>
<thead>
<tr>
<th></th>
<th>Gold Standard = Video Step Data + FSR Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heel Step Occurred</td>
<td></td>
</tr>
<tr>
<td>True Positive</td>
<td></td>
</tr>
<tr>
<td>False Negative</td>
<td></td>
</tr>
<tr>
<td>Heel Step Did Not Occur</td>
<td></td>
</tr>
<tr>
<td>False Positive</td>
<td></td>
</tr>
<tr>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>
Using these definitions, the number of true positives, false negatives, and false positives were calculated by comparing synchronized video heel steps and algorithm detected heel steps (i.e. the times when a sound played). These values were used to calculate the system heel step detection accuracy and percentage of false positives. Video data were also reviewed to investigate reasons for decreased accuracy (i.e. undetected events) and increased error rates (i.e. false positive events). Of note, accelerometer, gyroscope and magnetometer data were not used in this study, but were collected for future reference.

User feedback from subjective tools (i.e. post-test questionnaires, interviews) were qualitatively assessed and described via frequency counts and descriptive statistics where appropriate. Additionally, to analyze data collected during the interface testing with therapists, the percentage of users who successfully navigated through each task was calculated. Audio recordings were transcribed and feedback was categorized into areas for improvement for the system via simple content analysis.

3.3 Results

3.3.1 Participants

Children

Four children (P1, P2, P3, P4) were recruited for the study between the ages of 4 to 5 years as summarized in Table 14. Three of the children had cerebral palsy (P1, P3, P4). One of the children was recovering from a left middle cerebral artery stroke (P2). Finally, one of the children was also autistic in addition to having hemiplegic cerebral palsy (P4). All four of the children used ankle-foot orthoses (AFOs) and each child wore them for both sessions.
Table 14: Child Participant Demographics. (CP=Cerebral Palsy; MCA=Middle Cerebral Artery; ASD=Autism Spectrum Disorder; P1, P2, P3, P4 represent the first, second, third, and fourth child participants, respectively)

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Age</th>
<th>Gender</th>
<th>Diagnosis</th>
<th>GMFCS Level</th>
<th>Session with Music</th>
<th>PT/PTA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>5</td>
<td>M</td>
<td>CP</td>
<td>2</td>
<td>2</td>
<td>T1</td>
</tr>
<tr>
<td>P2</td>
<td>5</td>
<td>M</td>
<td>Left MCA Stroke</td>
<td>1</td>
<td>1</td>
<td>T3</td>
</tr>
<tr>
<td>P3</td>
<td>4</td>
<td>F</td>
<td>CP</td>
<td>1</td>
<td>2</td>
<td>T3</td>
</tr>
<tr>
<td>P4</td>
<td>4</td>
<td>F</td>
<td>ASD and CP</td>
<td>1</td>
<td>1</td>
<td>T2/T4</td>
</tr>
</tbody>
</table>

**Therapists**

Four therapists were recruited as summarized in Figure 15. All of the therapists were females between the ages of 26-40 years of age. Two of the participants were PTs, while one of the participants was a dedicated PTA and the last participant was both a physical therapy and occupational therapy assistant. The majority of users (n=3) reported that they had an average level of computer proficiency, while one participant reported an advanced proficiency. The same therapist reported frequent computer usage while the majority (n=3) of therapists reported using a computer a few times a day. All therapists (n=4) reported that they rarely use a tablet or smartphone; however, T2 reported to the researcher that if she had one, she would use it all the time.
Figure 15: Participant demographics for Usability Testing with Therapists. (T1, T2, T3, and T4 represent the first, second, third, and fourth therapist participants, respectively)

3.3.2 Phase 1: System Design and Usability Testing

System Performance

The algorithm was found to be on average 88.3% (SD=6%) accurate in detecting heel steps and triggering sounds. The percentage of false positive detections of heel steps was 13% (SD=7%).

Child Interaction and Usability

The results of the dichotomous Yes/No questions of the child’s post-test questionnaire are tabulated in Table 15. In addition to these results, children reported that they would like to hear sounds other than the music such as car (P1) and bus sounds (P2), and that they would like to hear different songs such as “Wheels on the Bus” (P4). Children also reported their likes and dislikes of the Musical Steps system. They reported that they liked walking (P2), the sounds (P3), and the interaction with other people and toys (P4, therapist-reported). Children disliked
walking (P1), walking without the sounds (P3), and walking backwards (P4, therapist-reported). Only one child (P1) understood the question regarding how to make the Musical Steps system better and reported car noises would make it better.

Table 15: Summary of results for user preference evaluation tools

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child Post-Test Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>I liked the sounds.</td>
<td>Yes (n=4)</td>
</tr>
<tr>
<td>The sounds were too loud.</td>
<td>No (n=3), Yes (n=1)</td>
</tr>
<tr>
<td>The sensors were annoying to wear.</td>
<td>No (n=3), Yes (n=1)</td>
</tr>
</tbody>
</table>

Therapist Interaction and Usability

The summary of system usability results from therapists are tabulated in Table 16. All therapists (n=4) successfully navigated through the prototype clinician interface without any need for guidance. All therapists (n=4) agreed that the performance data provided by the system could be helpful. Therapists (n=3) disagreed that the benefit of the system did not justify the added setup time to use it. However, T3 reported a neutral position on whether the benefit of the system justifies the added setup time or not. She reported that she thought it depended on the specific use-case of the system. She thought that “to combine this activity along with other exercises might be time consuming to set up.” However, she also noted that if an entire therapy session is being devoted to gait training, then the “setup time + activity could be useful in a therapy session.” All therapists either disagreed (n=3) or strongly disagreed (T2) that the system was too difficult to use. Finally, all of the therapists (n=4) reported that they agreed both that they would use the system if it was available and that they thought it would be a useful tool for therapy sessions.

Table 16: Summary of results for usability evaluation tools

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usability Testing Task Completion Rate</strong></td>
<td></td>
</tr>
<tr>
<td>Start a session</td>
<td>Task successfully completed (n=4)</td>
</tr>
<tr>
<td>End a session</td>
<td></td>
</tr>
<tr>
<td>Review session summary page</td>
<td></td>
</tr>
<tr>
<td><strong>Therapist Post-Test Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>The performance data (e.g. number of steps taken, cadence) from the Musical Steps system could be helpful to me.</td>
<td>Agree (n=4)</td>
</tr>
<tr>
<td>Statement</td>
<td>Agree (n=4)</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>The benefit of the system does not justify the added setup time in using it.</td>
<td></td>
</tr>
<tr>
<td>The system is too difficult to use.</td>
<td></td>
</tr>
<tr>
<td>I would use this system in regular therapy sessions if it was available.</td>
<td></td>
</tr>
<tr>
<td>I think the Musical Steps system is a useful tool to use in therapy sessions.</td>
<td></td>
</tr>
</tbody>
</table>

With regards to the system hardware, therapists were also asked to report on how they would improve the system. T1 thought the sensors were “too clunky” on the back of the child’s shoes. T2 thought that the volume should have been louder in the sessions. She also thought that it would be beneficial to “offer visual aid” to children in conjunction with the musical feedback. She thought that this could be done through “lights flashing and sound to increase feedback to client”. Finally, T4 though the system should be customizable for how it rewards children. Specifically, she thought it would be beneficial to be able to control which sensor(s) (i.e. left, right, or both) play(s) music when heel contact occurs. She would also like to be able to know “how close [a client’s] heel is to the ground” and also be able to customize the system so that it could “play music when [the] client reaches a certain threshold of how close [his/her] heel is to [the] ground.” Finally, she thought there was potential in the data that could be provided by the system in terms of the system’s use-cases. She thought that the client could “use [the] sensors during treatment sessions or at home to see how close [his/her] heel is to the ground when walking and to see if this changes with different interventions” such as “AFO, modifications to AFO, physiotherapy, Botox”.

With respect to the therapist interface software specifically, therapists reported a number of desired improvements. For instance, after the first task of starting a session was completed, therapists reported that they would expect that the system should automatically connect the sensors (T2, T4). T2 stated that after pressing the ‘Begin Session!’ button, “I would want the sensors to be turned on…everything should be starting up, ready to go.” The information displayed on the session screen may be ambiguous in terms of number of steps displayed. T2 thought that the steps displayed were the total number of heel steps, which T4 thought that the steps displayed were the total number of all steps taken. The layout and organization of information across screens was clear to therapists (T1,T2,T3); T3 stated, “I think it’s pretty clear and organized. There’s not a lot of information, but it’s exactly what you’re looking for. So for me it’s clear and it’s good. Easy to read.” Therapists appreciated the confirmation
popover presented prior to ending the session (T2, T3, T4); T4 stated, “I think that’s good in case by accident you pressed it.” However, T3 specifically would like a save button or to see the word ‘Save’ on the screen (“…I probably would be worried that nothing got saved, so I would just need like a save button somewhere”).

Therapists also reported that they would like additional information displayed for sessions including speed (T2), step length (T1), stance time (T4), number of heel strikes for slow versus fast walking (T2, T3), average height of heel off the ground (T2), and split information for left versus right foot (T4). T4 explicitly emphasized that, “For a child like we had today, I would absolutely want to know the left versus the right because for a child with hemiplegia I really wouldn’t be too concerned with the one side, and so I would assume that that side would have 100% of heel strikes and that would skew the numbers if we only cared about the one side, so definitely want to know left versus right.” Therapists would also want to be able to calibrate the system for specific clients’ speeds by performing baseline measures of fast versus slow walking. (T2, T3, T4) Therapists mentioned that they would like to be able to see each client’s progress over time (T1, T2, T3) and would like to be able to store data for multiple clients (T1, T2).

Overall, therapists expressed their interest in the system and reported on their thoughts of how they would adopt it into their therapy programs. One therapist reported seeing herself using the system as an assessment tool both for performing baseline measures and periodic assessments throughout the year. (T3) Alternatively, one therapist suggested that the system could be used to assess changes pre- and post-treatment, but was unsure whether it would be useful for every single session. (T2)

Finally, therapists also reported on the type of interface that they would prefer to use for Musical Steps. Therapists had mixed feelings on which form factor would be ideal for therapy sessions. T1 did not report a preference, two therapists thought that a tablet would work (T2, T3), but T4 thought that tablets can get in the way and are not ideal in the treatment of young children that require much attention and guidance. She stated, “…if you’re with a young child and you’re running around all over the place you can’t be holding really anything because you have to have your hands free to be working with them and if it’s on a table somewhere it
can get stolen, it can fall off, someone could knock it off…so it’s a little tricky to have anything you have to hold onto.” T4 suggested instead that, “it’s not a ton of information, so for our normal work, it would never fit on a screen because it’s all kinds of documentation, but this is just a quick snapshot, I would say probably on some kind of smartphone…with a big enough screen but you could then have a zipped pocket and shove it back in….I would imagine that would be the easiest.”

3.3.3 Phase 2: Feasibility Studies

Motivation

Enjoyment/ Fun

The summary of enjoyment/fun results for children are tabulated in Table 17.

**Table 17: Summary of results for enjoyment/fun evaluation tools**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Smileyometer</strong></td>
<td></td>
</tr>
<tr>
<td>How much fun was it to do this session?</td>
<td>Average with Music: 4 &lt;br&gt; Average without Music: 3.5</td>
</tr>
<tr>
<td><strong>Therapist Post-Test Questionnaire</strong></td>
<td>The client enjoyed the Musical Steps system. &lt;br&gt; Agree (n=2), Strongly Agree (n=2)</td>
</tr>
<tr>
<td><strong>Blinded Video Analysis Questionnaire</strong></td>
<td>The child appeared to enjoy the therapy session. &lt;br&gt; Average with Music: 3.25 &lt;br&gt; Average without Music: 3.75</td>
</tr>
<tr>
<td><strong>Child Post-Test Questionnaire</strong></td>
<td>I liked walking with the sounds more than without. &lt;br&gt; Yes (n=4)</td>
</tr>
<tr>
<td></td>
<td>Do you want to make the sounds happen again? &lt;br&gt; Yes (n=3), No (n=1)</td>
</tr>
<tr>
<td></td>
<td>I had fun walking. &lt;br&gt; Yes (n=3), No (n=1)</td>
</tr>
</tbody>
</table>

Distraction

The summary of distraction results for children are tabulated in Table 18.

**Table 18: Summary of results for distraction evaluation tools**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Therapist Post-Test Questionnaire</strong></td>
<td></td>
</tr>
</tbody>
</table>
The sounds distracted the client and interfered with the therapy sessions in a negative way. Disagree (n=3), Neither Agree or Disagree (n=1)

**Blinded Video Analysis Questionnaire**

The child seemed distracted during the therapy session. Average with Music: 3.5 Average without Music: 2.75

### Engagement

The summary of engagement results for children are tabulated in Table 19.

**Table 19: Summary of results for engagement evaluation tools**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blinded Video Analysis Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>The child appeared engaged/motivated during the session.</td>
<td>Average with Music: 2.5 Average without Music: 3.5</td>
</tr>
</tbody>
</table>

### Feedback

#### Sense of Agency

The summary of the child’s sense of agency results for children are tabulated in Table 20. Notably, P4 did not understand the questions regarding how she made sounds play and whether the system made her want to use her heels so could not report on these.

**Table 20: Summary of results for sense of agency evaluation tools**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Therapist Post-Test Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>The client understood how the Musical Steps system worked and how to activate the sounds.</td>
<td>Agree (n=3), Disagree (n=1)</td>
</tr>
<tr>
<td>The sounds helped the client focus more on achieving heel strikes.</td>
<td>Neither Agree or Disagree (n=3), Agree (n=1)</td>
</tr>
<tr>
<td><strong>Child Post-Test Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>How did you make the sounds play?</td>
<td>Not sure (n=1), stepping (n=2)</td>
</tr>
<tr>
<td>The musical walking made me want to use my heels when I walked.</td>
<td>No (n=3)</td>
</tr>
</tbody>
</table>
Practice

Achievement of Therapeutic Goal

The summary of the child’s achievement of therapeutic goal results for are tabulated in Table 21.

**Table 21: Summary of results for achievement of therapeutic goal evaluation tools**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gait Metrics</strong></td>
<td></td>
</tr>
<tr>
<td>Percentage of total steps taken with heel contact</td>
<td>Average (and standard deviation) with Music: 73.8% (35.5%)</td>
</tr>
<tr>
<td></td>
<td>Average (and standard deviation) without Music: 79.1% (29.5%)</td>
</tr>
<tr>
<td><strong>Therapist Post-Test Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>The client performed better (e.g. achieved more heel strikes) with the musical feedback.</td>
<td>Agree (n=1), Neither Agree or Disagree (n=1), Disagree (n=1)</td>
</tr>
<tr>
<td><strong>Blinded Video Analysis Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>There was evidence that the use of heel strikes was encouraged.</td>
<td>Average with Music: 3</td>
</tr>
<tr>
<td></td>
<td>Average without Music: 3</td>
</tr>
</tbody>
</table>

Task Complexity/Difficulty

Exertion

The summary of children’s exertion are tabulated in Table 22.

**Table 22: Summary of results for exertion evaluation tools**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P-CERT</strong></td>
<td></td>
</tr>
<tr>
<td>How difficult was it to do this session?</td>
<td>Average with Music: 7.5</td>
</tr>
<tr>
<td></td>
<td>Average without Music: 6</td>
</tr>
<tr>
<td><strong>Therapist Post-Test Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>The client exerted themselves more when there was musical feedback.</td>
<td>Strongly Agree (n=2), Agree (n=1), Neither Agree or Disagree (n=1)</td>
</tr>
<tr>
<td><strong>Blinded Video Analysis Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>The child appeared to exert themselves to the best of their ability.</td>
<td>Average with Music: 3</td>
</tr>
<tr>
<td></td>
<td>Average without Music: 3.25</td>
</tr>
</tbody>
</table>

Fatigue

The summary of children’s fatigue results are tabulated in Table 23.
### Table 23: Summary of results for fatigue evaluation tools

<table>
<thead>
<tr>
<th>Tool</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blinded Video Analysis Questionnaire</strong></td>
<td></td>
</tr>
</tbody>
</table>
| The child was fatigued by the therapy session. | Average with Music: 3.25  
Average without Music: 2.75 |
| **Child Post-Test Questionnaire**  |                                                  |
| I felt tired.                     | Yes (n=4)                                        |

### 3.4 Discussion

#### 3.4.1 Key Findings

**Phase 1 Findings: System Design and Usability Testing**

1. **Technology**
   - The detection algorithm was found to be moderately accurate with an average of 88.3% (SD=6%) of heel steps detected and an average of 13% (SD=7%) of steps taken resulting in false positive triggers.

2. **Usability of Musical Steps for children:**
   - Children liked using the system and the sounds produced by the system, but would prefer more customization of sounds.

3. **Usability of Musical Steps for therapists**
   - Therapists could successfully navigate through the prototype interface and found all the information displayed to be useful, clear, and well-organized.
   - Therapists thought the Musical Steps system would be a useful tool for gait training sessions with toe-walking children.
   - Usability issues still exist with the system that should be addressed in future redesigns.

**Phase 2 Findings: Feasibility Studies for Musical Steps**

The following presents the key findings with respect to the feasibility of Musical Steps for eliciting factors of neuroplasticity:
1. Motivation

- Children found the system to be fun and enjoyed therapy sessions where they walked with the musical feedback.
- The sounds may have distracted the children from engaging in the tasks of the therapy sessions.

2. Feedback

- Children may not have been able to understand how to produce the musical feedback and, therefore, did not perceive the feedback to be linked to the therapeutic goal.

3. Practice of the therapeutic goal

- Children exhibited less heel contact and, thus, achieved less practice of the therapeutic goal, in sessions where there was musical feedback.

4. Task Complexity/Difficulty

- Children exerted themselves more and were more fatigued in sessions when there was musical feedback suggesting that the system has the potential for promoting physical activity during therapy sessions.

While strong conclusions cannot be drawn, it appears that the system may be beneficial for enhancing enjoyment/motivation as well as exertion during physical therapy sessions, however, the inability to link the musical feedback to the therapeutic goal (i.e. heel contact) is a critical flaw in the system and/or use of the system that would preclude processes of neuroplasticity. This aspect should be the priority focus in future reiterations.
3.4.2 Study Strengths & Limitations

Study Design Strengths

While the small sample size precluded any statistical comparisons, four participants is adequate for preliminary testing and is estimated to identify approximately 80% of usability issues [48]. Thus, more participants may not have been profitable in terms of identifying high priority usability issues, and therefore this may have actually been a study strength as it demonstrated that a large number of usability issues can be identified even with a small sample size.

Study Design Limitations

While this study aimed at providing an example study using the proposed VRT evaluation framework, the steps followed in this study did not precisely adhere to the ideal approach that was intended. Most notably, Phase 1 and Phase 2 were performed simultaneously in one round of testing given time constraints which did not allow for design reiterations to fix issues identified in Phase 1 prior to beginning Phase 2. It is strongly recommended that future studies adopting the evaluation framework should address usability and system design issues prior to evaluation of factors affecting neuroplasticity, as we hypothesize that those issues can be highly influential on the outcomes of motivation, feedback, practice, and task difficulty/complexity.

Additionally, this case study utilized a quasi-experimental, randomized AB study design which limits internal validity of conclusions drawn from the results. Internal validity would have been strengthened with greater number of participants or use of an ABA design. As such, we were unable to draw strong conclusions from this study outside of identifying usability issues. Of note, gait data was not collected for P4’s second session due to technical difficulties.

The main tools used were gait metrics, the Pictorial Children’s Effort Rating Table (P-CERT) (APPENDIX C), the Smileyometer (APPENDIX D), custom post-test questionnaires for both the child (APPENDIX E) and the therapist (APPENDIX F), a custom blinded video analysis questionnaire (APPENDIX G), and unstructured post-test and post-analysis interviews with the therapists. The P-CERT and Smileyometer have been validated previously [86] [88], but
not for the young age range of children who participated in this study. Unfortunately, no validated tools were used in the evaluation of this technology, nor do they exist for the child population of interest. However, these tools have been previously used in research for a similar population of children [89] with some success. Thus, they were adopted here in the hopes that they would not only provide feedback on the Musical Steps system, but also lead to further insights into the evaluation of rehabilitation technologies designed for young children, in general.

After reviewing the muted videos, the blinded observer also reported on the Musical Steps system and protocol used. She thought that the sessions were not very realistic in terms of the duration for which the participant was asked to perform a single therapy task (i.e. straight, overground walking). She thought that this may have greatly influenced the children’s enjoyment and engagement in the sessions as 10 minutes of walking (even with a break in between) may be too long for them to stay engaged. This duration was chosen for the protocol based on recommendations from a previous study [90]. However, she reported that a better strategy would be to include other activities (e.g. heel presses, step ups) in the sessions to make it more realistic in terms of how therapy sessions are normally performed. On the video analysis questionnaires, she indicated that the activity needed to be varied (P1-A, P2-B) and that interest, motivation, engagement, and/or enjoyment were an issue for both sessions with P2, P3, and P4. She also commented that it was very difficult to report on some of the questions in the questionnaire because of the lack of sound to hear what kinds of comments the children were making.

While offline learning [91] may be an important issue to consider, sessions were at most 1 week apart with no other therapy sessions in between; therefore, this was not expected to be an issue. Furthermore, sessions were randomized so any effect that this might have had would have been seen in both children who had sessions with music first and without music first.

Finally, as is the case for most research studies with children, children may have felt uncomfortable with being watched and studied as they walked in the sessions. While every attempt was made by the researcher to make the child feel as comfortable as possible, this may still have had an impact on how the child felt during the sessions and could have reduced the
potential enjoyment factor of hearing the sounds when walking. Future research should investigate methods of data collection that minimize this feeling of being watched in children, including little or no video recording and reduced number of researchers present during sessions to simulate as closely as possible, regular therapy sessions.

Protocol Deviations

Some protocol deviations were experienced during the sessions, including the length of the break and protocol used for P4.

Length of the Break

The length of the break was defined as the difference in time between the last step of the first circuit and the first step of the second circuit. During sessions, this time was approximated using a stopwatch. It was difficult to get children to take precisely the same amount of break time for both sessions, but the timing was approximated. This variability was likely negligible, but nevertheless, it may have introduced some variability in terms of exertion levels. It was also difficult to get children to actually rest/sit down during the break, especially P3 and P4.

P4

Several deviations were required in the case of P4 from the standard protocol. Firstly, the same therapist could not attend both sessions, so her PTA attended the first session with music, and her PT attended the second session. As noted, P4 was autistic thus had more trouble than her peers in following the specific instruction regarding fast/slow/normal and forward/backward walking. P4 especially had trouble understanding the concept of walking backward so very little data was acquired for this walking state, but what data was acquired for her walking backward occurred when the PTA and mother held her hands on both sides of her and walked backwards with her, or when the PT walked forwards holding P4 to guide her walking backwards. Adherence to protocol was difficult, since she had trouble concentrating and understanding what was being asked of her. Toys were used to motivate her to walk back and forth in the study area. While this may have affected her motivation and reward levels during the sessions, the toys were introduced in both sessions, so any confounding factors should have been negated across sessions. P4 also had significantly more trouble answering the
questionnaires on her own; so her mother and therapists helped her to understand and interpret her response, or they answered for her. Times when the therapists reported for her were noted. Finally, due to usability issues, P4’s second session could not be used for evaluating change in gait metrics or system accuracy.

**System Design Limitations**

Technology limitations were also experienced. The detection algorithm was only designed for straight, over-ground walking and was evaluated for its detection of heel steps only in this context. Data, including foot shuffling and turning around on the 10m walkway, were not included in the analysis of the algorithm’s ability to detect heel steps and as such, children may have been over-rewarded by sounds (i.e. false positive detections). This may have contributed to children’s confusion as to the cause of the sounds. The system was designed in this way because FSRs were not expected to be able to handle these conditions (e.g shuffling, changes in direction) on their own. For this level of detection, inertial sensors are likely needed. Inertial sensors were not used in this study as initial data with the population of interest was needed to train detection algorithms to accommodate pathological gait prior to testing with the system. Collection of these data was carried out during the usability studies for future development applications and is an additional contribution of this study.

The FSR threshold was set to its maximum possible sensitivity and the same threshold was used for all 4 participants. However, the threshold may have been too high for some children, depending on their weight, and how much they use their heels. Future work should include the ability for therapists to utilize baseline weight measurements for each child and automatically calculate the threshold for the child’s ideal heel strike or heel contact event. While this would have been ideal in this study, the threshold was met by all children during the sessions and therefore, was deemed acceptable for the participants and did not affect accuracy of the algorithm.

In addition, the placement of FSRs was approximate. Other studies have utilized pressure mapping analysis prior to performing gait study with FSRs, but these studies utilized smaller sensors than those used in this study and since the children of this case study were very young, their shoes were small and the sensors covered the majority of the metatarsal and heel areas of
their shoes. Therefore, approximate placement of the FSRs was deemed acceptable for this case study. Nevertheless, there is the possibility that the FSRs may not have been placed on the area of maximal pressure, and as a result this could have resulted in the sensors’ threshold not being met and steps not being detected by the algorithm.

Further, children were not explicitly told to wear the exact same shoes for every session. This may have introduced variability in terms of their ability to get their heel down, for example, if one pair of shoes was stiffer than the other. Children using assistive devices such as AFOs and/or heel plates wore these devices for both sessions.

Researcher-noted Usability Issues

It is important to note that therapists were only asked to interact with the system’s prototype interface on the computer. So while they reported on the whole system’s usability, they did not interact with the sensors or the custom software used in the gait sessions; this setup and maintenance was performed by the study researcher who noted some usability issues with the system before, during, and after the sessions. The issues are summarized in Table 24 and are categorized into system management, software, sensor physical connectivity, sensor attachment, and technology limitations issues.

**Table 24: Summary of system usability issues identified by the researcher**

<table>
<thead>
<tr>
<th>System Management Issues</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>System usage is not a streamline process</td>
<td>The system’s setup and maintenance process required the user to perform many tasks for each session. This placed cognitive load on the user. As such a checklist was necessary for the researcher to remember everything that needed to happen in preparation for sessions, during sessions, and after sessions. This is particularly important to note as use of the system would require therapists to perform all these steps before, during and after sessions if they were to adopt this technology.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Software Issues</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor Bluetooth® connection to software is not reliable</td>
<td>The current software does not automatically connect the sensors in a simple, reliable manner. As such, the researcher was forced to manually link the sensors over Bluetooth and input the computer’s assigned communication channels into the software.</td>
</tr>
<tr>
<td>System notifications are not provided to the user</td>
<td>During P4’s second session, the right Shimmer unit did not connect properly to the computer over Bluetooth® at the beginning of the data collection session. No</td>
</tr>
</tbody>
</table>
if the Shimmers do not connect properly | system error notifications were provided. As such, the system did not collect any data for this foot during the session. This error was not detected until after the session had been completed and data was being reviewed by the researcher.

System input commands are not intuitive | The custom software application was windows command prompt-based and relied on the memorization of keyboard shortcuts/character input commands. No graphical user interface (GUI) was available for graphical interaction with the commands.

**Sensor Physical Connectivity Issues**

| FSRs may detach from the connectors | During P1 and P4’s second session, one of the FSRs detached from the custom connectors. The researcher was forced to visually notice this and reattach them. No notification was provided by the software. |
| FSRs may tear during use | During P4’s second session, the left heel FSR was completely torn apart between its sensing area and the wire connector. No notification was provided by the software. |
| FSR connection to the Expansion Board was not sufficiently durable | During P1’s second session, the soldered connections between the expansion board and the FSRs were not durable enough to withstand the session, and the signal was attenuated, but not lost. After this incident, the connection method was redesigned to be much more durable and this problem did not occur again in the rest of the sessions. |
| Expansion Boards may disconnect from the Shimmer units | In P1’s second session, one of the expansion boards disconnected from its Shimmer unit. The researcher was forced to visually notice this and reattach it. No notification was provided by the software. After this incident, the connection method was redesigned to be much more durable and this problem did not occur again in the rest of the sessions. |

**Sensor Attachment Issues**

| Shimmer units may detach from the child’s shoes | The Shimmer units were attached to the child’s shoes using tape, a method previously used in [89]. However, due to increased mobility of children in gait training compared to stair climbing, children exhibited much more demand on the sensor attachment (e.g. running around, stomping their feet to make sounds, knocking their heels together during ambulation, etc.) Tape is not a sufficiently durable means of attaching the Shimmer units for these conditions. As such, the attachment to the child’s shoes needed to be fixed in 6 out of 8 of the usability testing sessions with children. (P1, P2, P3, P4) |
| FSRs may detach from the child’s shoes | In P4’s second session, the toe FSR for the left foot repeatedly detached from the child’s shoe. |
| Sensor attachment process was | In total, there were 6 sensors that needed to be attached to each child for each session. While this process was optimized to be as streamline as possible, it still |
cumbersome and time-consuming took quite a long time (~5 minutes) for the sensors to be attached. It was quite difficult to get children to sit still for this long (especially P3 and P4).

Limited guidance was provided for distinguishing the left and right foot sensors

Since the foot assignment for each Shimmer changed every time they were connected, the only indicators for distinguishing the left and right foot sensors were green and yellow/orange LEDs on the Shimmer units. The green LED indicated the right foot, while the yellow/orange LED indicated left. While a simple mnemonic could be used (yeLlow=Left; gReen=Right), this presents an additional cognitive load on the user.

<table>
<thead>
<tr>
<th>Technology Limitations</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery life limitations of the laptop were not conducive to mobility requirements of gait therapy sessions</td>
<td>Therapists require an interface that is mobile and capable of having a battery that is capable of powering the system pre-, during, and post-therapy sessions. The laptop used in this case study was relatively older and discharged battery power very quickly. The researcher was very cognizant of battery life of the laptop during sessions, and was required to conserve power as much as possible if the laptop was idle (e.g. participants were late for the session).</td>
</tr>
<tr>
<td>Bluetooth® range limits the distance participants can be from the computer</td>
<td>Due to the limited connection strength of Bluetooth®, children must always stay relatively close to the data acquisition unit (in this case, the laptop). Some connectivity issues were experienced when children wandered too far away from the research environment.</td>
</tr>
<tr>
<td>Devices required charging after each session</td>
<td>The laptop and sensors relied on limited battery life for each session and as such required charging after each session.</td>
</tr>
</tbody>
</table>

In general, the researcher noted that there were insufficient error messages for sensor malfunctions. Additionally, the user’s system state awareness is inadequate. For example, the user may be unsure at any point in time during the system’s use whether or not the sensors are currently connected and are collecting data, or whether the sounds are muted or not.

3.4.3 Future Work

System redesign is necessary to address the usability issues

In order to proceed in the evaluation of the Musical Steps system, redesign is needed to address the various usability issues identified during testing. Redesign includes high, medium, and low priority recommendations. High priority recommendations identified include the need for (i) a higher accuracy and more robust detection algorithm, (ii) a fully functional clinician interface,
(iii) a more reliable method of connecting the sensors over Bluetooth®, and (iv) a more robust means of attaching the sensors to children’s shoes.

i. Design a higher accuracy and more robust detection algorithm

   o It is disadvantageous to use FSRs for therapy sessions due to their cost, fragility, and setup time. FSRs were replaced for every study session. While this ensured maximum quality of data collected, it would be impractical to replace them for every therapy session if this technology were to be adopted into regular gait therapy programs. Additionally, the FSRs were somewhat time-consuming to attach to the children's shoes. It is recommended that FSRs should not be utilized in future iterations as there were many usability issues associated with sensors and their attachment to the Shimmer units.

   o It is hypothesized that accelerometer and gyroscopes will be able to offer more information regarding the biomechanics of gait (e.g. toe-heel gait patterns) than using FSRs alone. Many studies have used accelerometers as a means of gathering relevant gait metrics. [74] Accelerometry, while not a new technology, is relatively new in terms of its applications to the study of pathological gait, including children with CP. While there is considerable literature available for accelerometry use with typically developed children, the literature is very limited when used with children with pathological gait and even less for toe-walking children. Previous research has demonstrated the use of accelerometry in a stair-climbing musical feedback system with some success. [89]

   o While Musical Steps was designed to rely only on FSRs to trigger automated musical feedback, it would be ideal to instead use only the Shimmer 9DoF sensing units. The accelerometer, gyroscope, and magnetometer data collected during this study should be utilized to develop and test a novel method of detecting heel strikes in toe-walking children.

ii. Integrate custom software functionality into the prototype

   o A functional user interface should also be integrated into the system. While a prototype user interface was designed and tested in this study using a laptop, further study should go into the integration of this user interface on various
platforms (ex. tablets, smartphones) to investigate the interaction between PTAs and the technology to review gait metrics, choose feedback types (ex. songs, animal noises), control conditions for triggering sounds, adjust sound volume during gait therapy sessions, and track a child’s progress over time. All functionality available through the custom Microsoft Visual C++ software application used for the study should be transferred into the user interface. Transferring these functions into a GUI should address the issues surrounding streamline process of system setup, intuitiveness of input commands, system notifications and error prevention.

iii. Address Bluetooth connectivity issues and integrate a reliable, automatic sensor connection method
   - Future work should fully investigate the root cause of sensor wireless connectivity issues especially prior to its introduction into actual therapy sessions or home use. The process for connecting the sensors to the data acquisition device should be seamless, intuitive, and provide clear error notification messages if the sensors have not been connected properly.

iv. Redesign attachment method of sensors to children’s shoes
   - The need for reattachment of sensors during sessions caused interruptions in the child’s attention and could have contributed to the children’s distraction and enjoyment during the study sessions. During usability testing, one of the therapists commented on the setup time involved with the sensors being a problem if they were to be used in a session where other therapies were being performed. Another therapist was dissatisfied by the sensors falling off of the child’s shoes repeatedly during the session.
   - For future research, a better method of attachment would be useful and more conducive to adoption into therapy sessions. It is hypothesized that if the method of attachment was simpler, therapists would be more willing to quickly attach them during any session to get data/produce feedback for the children and would be more satisfied with the system. Therefore, future work should work on improving the method of sensor attachment to children’s shoes.
A medium priority recommendation identified was the need for additional gait metrics to be tracked and displayed by the system.

- Additional metrics (i.e. speed, step length, stance time, and average height of heel from ground) should be investigated and, if possible, added to the functionality of the system.

Finally, low priority recommendation identified was the need (i) to be able to track multiple clients’ data, and (ii) for the system to provide simple guidance and instructions for distinguishing left foot versus right foot sensors prior to attachment to children’s feet.

**System redesign is necessary to address the feasibility issues**

The critical issue identified by this study is the lack of motivation for the child to achieve the therapeutic goal of heel contact. We hypothesize that this was due to the child’s inability to make a clear link between the musical feedback and their use of heel contact when taking steps. This may be due to the young age of the children and their inability to monitor their own behaviour (i.e. how frequently they were making sounds) and understand their cause and effect relationship with the system. To remove possible confounding factors, our protocol excluded verbal cues from the therapist. The result of this lack of coaching may indicate that the child-therapist interaction is very important to the success of this VRT technology. In [89], verbal cues were utilized alongside musical feedback and resulted in positive effects on promoting the therapeutic goal of reciprocal stair-climbing steps. We hypothesize that since the protocol for the Musical Steps study specifically asked therapists not to promote heel contact, this broke the link between the child and the therapist necessary for the child to fully benefit from this system. This may further strengthen the model we developed in the proposed framework’s first phase. Therefore, future work should investigate the effect of the addition of therapist cues provided in conjunction with the Musical Steps system on the motivation of the therapeutic goal in toe-walking children.

One of the main motor learning principles of motor learning theory is that there are stages of learning. The three stages of learning are (i) cognitive, (ii) associative, and (iii) autonomous. [1] Of particular interest are the first 2 stages where therapists are required during the practice of the motor task. During the cognitive stage, the movement required for the task may not be
obvious to the child and thus, performance is heavily dependent on the child’s conscious effort, which is normally encouraged through a therapist’s verbalization of the movement. Once children become more comfortable with the movement required for the task, therapists will provide fewer cues and allow the child to independently adjust their movements; this is known as the associative stage of learning. [2] For this study, we assumed that the feedback provided by the system would adequately provide children with the cause and effect relationship between the heel contact and the music. Thus, we assumed that we could supplant the therapists’ verbalization of the movement with musical biofeedback, essentially jumping straight to the associative stage of learning. However, given responses on the feedback survey, children clearly did not understand the association. Therefore, future studies should factor in these stages of learning and allow for therapists to coach children in the early stages of presentation of a VRT, and gradually reduce therapist involvement over time once children become familiar with the VRT. While great effort should be placed on designing VRTs that are intuitive for children, training with therapists may be an integral part of the motor learning process in pediatric rehabilitation.

Also, rewards were designed to be given whenever children made heel contact with the ground for the first time in a given step. This meant that the goal of the study was to increase their heel contact overall, not specifically their heel strike. Further research could look at various other reward conditions. For example, children with hemiplegic cerebral palsy may tend to toe-walk only on one foot. In these cases, it desirable to only reward the affected side as opposed to both. The ability to customize reward conditions was reported by clinicians to be a desirable feature that would be very useful if Musical Steps was to be adopted into regular physical therapy.

Future studies with this system should also integrate other therapeutic activities to reduce the amount of time children are asked to perform straight walking, as this may decrease their engagement in the sessions. Therapy sessions with toe-walking children should be recorded using the sensors during the entire duration of the session to gather data on all activities that are performed and understand the various goals associated with each activity. From this data, algorithms for each activity or, if possible, one algorithm should be developed to account for all these activities and their goals.
3.4.4 Clinical Implications

Since children who toe-walk generally may experience a diminished endurance and distance travelled due to the inefficiency of their gait pattern [58], it is especially interesting to note that the Musical Steps system may encourage children to exert themselves more during therapy sessions, which could affect their overall endurance with continued use of the system. Perhaps, the system could even be provided to children and their families for use at home to both (a) promote physical activity, and (b) gather gait information between physical therapy sessions or over the course of an intervention for therapists to review from activities of daily life and outside of the context of physical therapy.

Therapists have indicated that they are unsure as to the practical aspects of this VRT technology as a specific intervention for treating toe-walking; this is primarily due to its preliminary nature. Within the framework that we developed, once the technology sufficiently iterates through usability testing and neuroplasticity feasibility testing, future work should focus on providing adequate evidence for functional outcomes that may be linked to this VRT technology within the broader context of rehabilitation in children. However, in the infancy of this technology, we have already identified its potential usefulness as a means for therapists to gather gait metrics especially within the context of periodic progress checking for children as part of their standard physical therapy, or as a simple means of gathering efficacy of an intervention for individual clients by measuring simple gait metrics pre-and post-typical interventions used for these children.

3.5 Conclusion

A novel VRT system was developed and evaluated preliminarily for its efficacy for promoting heel contact for toe-walking children. It demonstrated good system accuracy and promising results with respect to usability. However, future work should address the issues identified and move toward an accelerometry-based approach. This case study demonstrated the use of the novel evaluation framework for the evaluation of the novel ‘Musical Steps’ VRT technology.
Chapter 4

4 Conclusion

4.1 Overview

In summary, VRT technologies are being used increasingly to augment pediatric rehabilitation therapies. VRT technologies are of particular interest for their potential to create an environment rich in motivation and feedback that encourages practice of tasks with increasing complexity, as needed for effective neuroplasticity and motor learning. It is important to understand the extent to which VRT technologies are effective in eliciting these factors in neuroplasticity in order to better understand the “active ingredients” in VRTs and the value of VRT technologies. This thesis explored the research question: How should VRT technologies be evaluated? A scoping literature review was conducted to investigate current practices in VRT technology evaluation. Of the 21 studies included in this review, 67% included measures related to motivation, 38% measured the extent of practice/repetition of task, 29% assessed factors associated with task complexity/difficulty, and 5% evaluated interpretation or experience of feedback. While all of the studies reported functional outcome measures, it was evident that the ability to understand the mechanisms underlying the success or failure of VRTs is limited. In response, recommendations were made towards the development of a standard toolkit for evaluating the potential of VRT technologies for motor learning, and a novel, 3-phase evaluation framework was proposed. To demonstrate the viability of the proposed framework, a ‘Musical Steps’ system was created that uses force sensing resistors and a novel algorithm to reward heel contact with musical sounds for toe-walking children. Phase 1 (usability testing) and Phase 2 (feasibility testing) of the evaluation framework was carried out with 4 therapists and 4 children (aged 4 to 5 years). The system was 88% (SD=6%) accurate overall with a false positive detection rate of 13% (SD=7%). Initial usability tests (phase 1) were positive, but feasibility testing (phase 2) indicated that while children enjoyed the musical feedback, they did not understand the link between the sounds and the therapeutic goal. As such, increased heel contact was not observed and further reiteration of the technology is required. Detailed recommendations for future development of ‘Musical Steps’ were provided based on feedback from multiple perspectives (i.e. children, therapists, a blinded observer, researcher). This thesis conceptualized and demonstrated the value of a 3-phase framework to
evaluate VRT technologies with respect to their potential for eliciting neuroplasticity and motor learning. Use of this iterative, user-centered framework may lead to more efficient design of VRT technologies and a better understanding of the “active ingredients” in VRT interventions.

4.2 Summary of Contributions

This research thesis has provided the following contributions to the field of pediatric rehabilitation:

1. A comprehensive list of measures and tools that have been previously used in the evaluation of VRT technologies was developed. Recommendations were provided towards the development of a standardized evaluation toolkit for evaluating “active ingredients” of VRTs.

2. A novel 3-staged framework was developed based on the widely-accepted ideology and methodology of iterative and user-centered design. It was developed for the specific context of evaluating VRT technologies’ abilities to elicit motor learning in children with disabilities.

3. A novel ‘Musical Steps’ VRT technology was designed and tested for promoting heel contact in toe-walking children which, preliminarily, demonstrated good system accuracy and usability based on results of usability testing with child and therapist users.

4. A case study has been provided to act as a demonstration for the use of the novel evaluation framework that was developed for the evaluation of the ‘Musical Steps’ VRT technology. The system was evaluated for the first 2 (the most novel) phases of the framework.
4.3 Directions for Future Work

This research thesis presented a novel evaluation VRT technology framework for the context of motor rehabilitation. While preliminary conclusions have been made, future work is necessary to further develop this framework, namely, through the development of a standardized toolkit. Recommendations for future work toward the development and use of this standard toolkit are as follows:

- **More comprehensive evaluations are needed.** Studies should investigate all 4 aspects of neuroplasticity. New tools are needed particularly for evaluating feedback and characterizing task complexity.

- **More standard evaluation tools are needed.** Standardized questionnaires should be developed collaboratively that address the key factors of neuroplasticity and, also, aspects of system usability and user interactions.

- **Use a balanced mixture of objective and subjective measurement tools.** Subjective measures are essential for capturing users’ perceptions and interpretations. Objective measures can be used to corroborate subjective measures and can also reduce the number of questionnaires that must be completed.

- **Conduct multi-perspective evaluations.** Understanding interactions of each of the user groups (e.g. patients, therapists, parents) with the VRT technology as well as interactions between key stakeholders (e.g. therapist-child dynamics) in the use of these technologies is important to the overall evaluation of these systems.

- **Conduct context specific evaluations.** VRT technologies may be intended for use in the clinic, school, home, and/or in multiple contexts. The importance of evaluating VRT technologies within their intended context should be emphasized.
This research also presented the novel ‘Musical Steps’ VRT technology. A number of limitations were noted in this study as detailed in Chapter 3. Future work should aim to address these limitations. In summary:

- **System redesign is necessary to address the usability issues.** High priority recommendations should be addressed first. First, a higher accuracy and more robust algorithm should be developed using inertial sensors to compensate for the variability in the pathological gait of children. Second, custom software functionality should be integrated into the prototype. Third, Bluetooth® connectivity issues should be addressed and a reliable, automatic sensor connection method integrated in the system. Fourth, the attachment method of the sensors to the children’s shoes should be redesigned to be more robust.

- **System redesign is necessary to address the feasibility issues.** Future work should investigate the effect of the addition of therapist cues provided in conjunction with the Musical Steps system on the motivation of the therapeutic goal in toe-walking children. Further research could look at various other reward conditions. For example, children with hemiplegic cerebral palsy may tend to toe-walk only on one foot and hence, it may be important to limit rewards to the foot of interest.

### 4.4 Closing Remarks

Physical therapies can be difficult and demotivating to children. In strong contrast, technologies like the Nintendo® Wii are getting children to actively seek out opportunities for physical activity through fun games that allow them to explore new possibilities and experiences in the comfort of their own home. It is easy to see how desirable this is when you think of this kind of technology being used in the context of promoting therapeutic goals in children with disabilities. The idea that one can turn therapy into a game or play time is indeed an exciting prospect.
Even though the Musical Steps system did not specifically end up promoting heel contact in children, I still witnessed wonder in children’s eyes right from the moment I turned on the sounds in sessions as they began to realize that they were creating their own musical soundscape. When I saw them running around and stomping their feet, trying to figure out all the ways they could make the sounds play, I knew they had enthusiastically accepted the sounds and that their curiosity and playfulness had definitely been sparked.

Based on just a few experiences, I can personally attest to the potential and power of a simple VRT technology that augmented gait therapy into an opportunity for play in children. Based on the framework that was developed, we were able to understand why the technology did not specifically work at promoting a specific therapeutic goal. This study alone demonstrates the definite need for specificity in evaluation of these technologies, especially as therapies and goals become increasingly more complicated.

It is hoped that the contributions of this thesis will help to further identify VRT technologies that effectively promote therapeutic goals in therapy while also creating an environment full of wonder and magic for children to play and accomplish things they never thought possible.
References


treat impaired upper extremity motor function in children with cerebral palsy,"


[38] J. Howcroft, S. Klejman, D. Fehlings, V. Wright, K. Zabjek, J. Andrysek and E. Biddiss, "Active Video Game Play in Children With Cerebral Palsy: Potential for Physical Activity Promotion and Rehabilitation Therapies".


Appendices

A. Child Participant Demographic Questionnaire

<table>
<thead>
<tr>
<th>Participant Data Collection Form</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Section 1 (To be completed by the researcher)</strong></td>
</tr>
<tr>
<td>1. Participant ID</td>
</tr>
<tr>
<td>2. Session with Musical Feedback</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>3. Session without Musical Feedback</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Section 2 (To be completed by the Therapist)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Physical Therapist/Physical Therapy Assistant</td>
</tr>
<tr>
<td>5. Client’s Age</td>
</tr>
<tr>
<td>6. Client’s Gender</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>7. Client’s Diagnosis</td>
</tr>
<tr>
<td>8. Client’s GMFCS Level</td>
</tr>
<tr>
<td>GMFCS I</td>
</tr>
<tr>
<td>GMFCS II</td>
</tr>
<tr>
<td>GMFCS III</td>
</tr>
</tbody>
</table>
B. Therapist Participant Demographic Questionnaire

Therapist Demographic Survey

1. Participant Number

2. In what age range do you belong?
   - under 25 years
   - 26-40 years
   - 41-65 years
   - over 65 years

3. What best describes your professional role?
   - Physical Therapist
   - Physical Therapy Assistant
   - Other (please specify)

4. How would you rate your computer proficiency?
   - Expert - I am extremely proficient in using a wide variety of computer technologies
   - Advanced - I have acquired the ability to competently use a broad spectrum of computer technologies
   - Average - I demonstrate a general competency in a number of computer applications
   - Beginner - I am able to perform basic functions in a limited number of computer applications
   - Unfamiliar - I have no experience with computer technologies

5. How often do you use a computer during your work?
   - All the time
   - A few times a day
   - A few times a week
   - Very rarely

6. How often do you use a tablet and/or smartphone in your work?
   - All the time
   - A few times a day
   - A few times a week
   - Very rarely
C. P-CERT

How difficult was it to do this session?

D. Smileyometer

How much fun was it to do this session?
E. Child Post-Test Questionnaire

Please help us improve the musical gait system by answering these quick questions!

1. I had fun walking. □ Yes □ No
2. I felt tired. □ Yes □ No
3. How did you make the sounds play?
4. The musical walking made me want to use my heels when I walked. □ Yes □ No
5. I liked the sounds. □ Yes □ No
6. The sounds were too loud. □ Yes □ No
7. The sensors were annoying to wear. □ Yes □ No
8. I liked walking with the sounds more than without. □ Yes □ No
9. Do you want to make the sounds happen again? □ Yes □ No
10. Are there any other sounds you would like to hear?
11. What did you like best about Musical Steps?
12. What didn’t you like about Musical Steps?
13. How could make Musical Steps better?
F. Therapist Post-Test Questionnaire

1. The client exerted themselves more when there was musical feedback.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree or Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. The client performed better (e.g. achieved more heel strikes) with the musical feedback.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree or Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. The client enjoyed the Musical Steps system.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree or Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. The client understood how the Musical Steps system worked and how to activate the sounds.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree or Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. The performance data (e.g. number of steps taken, cadence) from the Musical Steps system could be helpful to me.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree or Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. The benefit of the system does not justify the added setup time in using it.
<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree or Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. The system is too difficult to use.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. The sounds distracted the client and interfered with the therapy sessions in a negative way.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. The sounds helped the client focus more on achieving heel strikes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. I would use this system in regular therapy sessions if it was available.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. I think the Musical Steps system is a useful tool to use in therapy sessions.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. I would improve the system by...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
G. Blinded Video Analysis Questionnaire

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neither Agree or Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The child appeared engaged/motivated during the session</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The child appeared to exert themselves to the best of their ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The child appeared to enjoy the therapy session</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The child seemed distracted during the therapy session</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>There was evidence that the use of heel strikes was encouraged.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The child was fatigued by the therapy session</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Additional comments regarding the therapy session:

__________________________

__________________________