MEASUREMENT OF INTERVENTION FIDELITY WITHIN PAEDIATRIC REHABILITATION

MEASUREMENT OF INTERVENTION FIDELITY WITHIN PAEDIATRIC REHABILITATION FOR CHILDREN WITH PHYSICAL DISABILITIES

By

Briano Marco Di Rezze, B.Sc. (Hon.), M.Sc. (OT), OT Reg (Ont.)

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfilment of the Requirements

for the Degree of

Doctor of Philosophy

McMaster University

Doctor of Philosophy (2012) McMaster University

(Rehabilitation Science) Hamilton, Ontario

TITLE: **MEASUREMENT OF INTERVENTION FIDELITY WITHIN PAEDIATRIC REHABILITATION FOR CHILDREN WITH PHYSICAL DISABILITIES**

AUTHOR: Briano Marco Di Rezze, B.Sc. (Hon.) (University of Toronto), M.Sc. (OT) (McMaster University)

Supervisor: Dr. Mary Law, Ph.D.

NUMBER OF PAGES: 240

Abstract

Measurement of intervention fidelity within paediatric rehabilitation for children with physical disabilities

Intervention fidelity examines the degree to which an intervention is delivered as planned. Generic fidelity measures incorporate the active ingredients of more than one intervention and characteristics common to all interventions. Three studies were conducted to define the active ingredients of intervention for children with physical disabilities and generate a generic fidelity measure. These studies involved: (1) describing generic fidelity measures; (2) generating essential attributes of paediatric rehabilitation; and (3) differentiating between two interventions to consistently rate the behaviours of the therapist, child and parent.

(1) In a narrative review of generic fidelity measures, five measures were identified within the psychotherapy literature. These measures presented a variety of approaches to examine fidelity, described psychometric property standards, and highlighted 37 non-specific intervention items that were relevant to paediatric rehabilitation.

(2) A consensus process with eight experts and interviews with seventeen clinicians working with children with physical disabilities generated 35 attributes that highlighted the general observed therapist and client behaviours essential within a successful intervention session.

(3) The Paediatric Rehabilitation Observational measure of Fidelity (PROF) was developed (30 items) to evaluate specific and non-specific behaviours within two occupational therapy and physiotherapy interventions for children with cerebral palsy. Six trained raters examined 25 intervention videos for psychometric testing. Results indicated that the PROF demonstrated good to excellent Inter-rater reliability and early construct validity.

These studies present an important starting point to observe and measure the active ingredients within paediatric rehabilitation, incorporating its dynamic nature involving the child and parent within the therapy process.

Acknowledgements

I want to thank Mary Law, my supervisor, for her endless amount of patience and encouragement throughout my PhD. Even though her wisdom was instrumental in helping me to navigate through this training, her partnership has made every step in the process enjoyable. It truly has been a pleasure working with her and my learning experience has been such a privilege.

I also wanted to thank Kevin Eva, Nancy Pollock and Jan Willem Gorter. As members of my thesis committee, they provided thoughtful insight and support throughout the program in such a comfortable and collaborative manner. They have each left an impression on me in terms of their perspectives and high standards of professionalism, which I will carry with me throughout my career. Kevin Eva was especially supportive and went way beyond the "call of duty" in providing thesis and career guidance during my PhD program, for this, I am truly grateful.

I am also thankful for the help that I received from the staff at *CanChild* Centre for Childhood Disability Research (especially Barb Galuppi) and the School of Rehabilitation Science (especially Stacy Maskell) for providing me with support to access materials and resources/equipment to conduct my research.

This work is dedicated to my entire family because none of this work would have been possible without your collective contributions – it really has taken a community to help me complete this PhD:

My in-laws (Linda and Tony) who have unconditionally supported my efforts from the beginning and have always found ways to help me to find time in the day to

work on this PhD. A special thanks to Anthony and Marco for not judging their zio for still being in school and not working like the rest of the adults!

My parents (Anna and Gino) who have always encouraged me to strive to do my best in any of my life goals, and who have contributed in so many ways to helping me reach these goals, including this PhD. They are great moral and ethical role models that have shaped my approach to research and the relationships I have built during this PhD.

My wife Mariana who has always encouraged me to pursue things that I am passionate about - regardless of how ridiculous it might sound. She is an extraordinary woman who has been the breadwinner, the caregiver, the glue connecting our family, and my voice of reason during the peaks and troughs of this PhD experience. I truly could never have completed this work without her. When we started this PhD journey together, it was just her and I, but along the way came our two precious children, Christian and Evangeline. Not only have my children taught me how to be more efficient with my time, but their arrival provided inspiration and new energy to help me progress through this PhD.

I would also like to acknowledge the funding support that I received throughout my graduate studies: the Canadian Institutes of Health Research (CIHR) Quality of Life Training Program in Rehabilitation Research; Autism Research Training (ART) Program Fellowship (CIHR); Ontario Graduate Scholarship; Children's Rehabilitation Research Network (CRRN) Graduate Student Scholarship; and *CanChild* Centre for Childhood Disability Research Studentship Award.

Preface

This thesis is comprised of three studies, each presented in an expanded version of what will be submitted as a manuscript for publication, except for Chapter 2, which is already accepted for publication.

For all three studies:

Briano Di Rezze defined the research questions, developed initial study designs, implemented the data collection process, secured ethical approval (where applicable) through the Faculty of Health Sciences Research Ethics Board at McMaster University, collaborated with personnel to access secondary data (for Chapter 4), conducted the analyses, and prepared the manuscripts.

Dr. Law, assisted with the refinement of the research questions, provided recommendations on the proposed methods, contributed to data collection process where applicable (i.e., agreement process in Chapter 3), provided her interpretation of findings, and provided ongoing editorial assistance during manuscript preparation.

When these studies were conducted the co-authors held the following roles: Dr. Mary Law, Ph.D., co-author, was Professor in the School of Rehabilitation Science at McMaster University; Dr. Kevin Eva, Ph.D., co-author, was Associate Professor in the Department of Medicine at the University of British Columbia; Dr. Jan Willem Gorter, MD, Ph.D., co-author, was Associate Professor in the Department of Pediatrics at McMaster University; and Prof. Nancy Pollock, O.T., M.Sc., co-author, was Associate Clinical Professor in the School of Rehabilitation Science at McMaster University.

Chapter 2

This chapter contains a manuscript entitled “A narrative review of generic intervention fidelity measures”. The authors are: B. Di Rezze, M. Law, J.W. Gorter, K. Eva, and N. Pollock. The design, data collection, data analysis and writing were conducted from December 2009 to January 2011. The manuscript has been accepted on May 9, 2012, for publication in the Journal of Physical and Occupational Therapy in Pediatrics.

Dr. Jan Willem Gorter contributed to the refinement of the research question and study design, contributed to the interpretation of findings, and provided editorial comments regarding the manuscript and its overall structure.

Dr. Kevin Eva contributed to the refinement of the research question and study design, contributed to the interpretation of findings, and provided editorial comments regarding the manuscript.

Prof. Nancy Pollock contributed to the refinement of the research question, contributed to the interpretation of findings, and provided editorial comments regarding the manuscript.

Chapter 3

This chapter contains a manuscript entitled: “General therapy behaviours in paediatric rehabilitation: essential attributes for intervention with children with physical disabilities”. The authors are: B. Di Rezze, M. Law, K. Eva, N. Pollock, and J.W. Gorter.

The design, data collection, data analysis and writing were conducted from January 2010 to May 2011.

Dr. Kevin Eva contributed to the refinement of the research question and study design, contributed to the interpretation of findings, and provided editorial comments regarding the manuscript and its overall structure.

Prof. Nancy Pollock contributed to the refinement of the research question, contributed to the data collection and interpretation of findings, as well as provided editorial comments regarding the manuscript and its clinical relevance to the field.

Dr. Jan Willem Gorter contributed to the refinement of the research question, contributed to the data collection and interpretation of findings, as well as provided editorial comments regarding the manuscript.

Chapter 4

This chapter contains a manuscript entitled "Development of a generic fidelity measure for rehabilitation intervention research for children with physical disabilities". The authors are: B. Di Rezze, M. Law, K. Eva, N. Pollock, and J.W. Gorter. The design, data collection, data analysis and writing were conducted from January 2011 to February 2012.

Dr. Kevin Eva contributed to the refinement of the research question, statistical analyses and procedures, interpretation of findings, as well as provided editorial comments regarding the manuscript and key points informing the future directions of this work.

Prof. Nancy Pollock contributed to the refinement of the research question, contributed to the pre-testing process of the measure and interpretation of findings, as well as provided editorial comments regarding the manuscript and its clinical relevance to the field.

Dr. Jan Willem Gorter contributed to the refinement of the research question, contributed to the pre-testing process of the measure and interpretation of findings, as well as provided conceptual editorial considerations and comments regarding the manuscript.

Table of Contents

List of Tables

List of Figures:

Appendices:

Chapter One: Introduction

Paediatric rehabilitation can be defined as, "a collaborative disciplinary endeavor designed to support children with disabilities in their efforts to achieve their highest potential, consistent with degree of impairment, given an inability or limitation in performing socially defined activities or roles" (Allen, Wilczynski, & Evans, 1997, p.27). This definition identifies the complexity of paediatric rehabilitation based on the involvement of a range of health disciplines, type of disability, and ability of the child and contextual and social demands. Furthermore, health care providers in paediatric rehabilitation need to be cognizant of the developmental level, preferences, roles, strengths and challenges of the child (Rodger & Keen, 2010). Another layer of complexity involves the perspective of the family because best practice in paediatric rehabilitation integrates the family within the intervention process (Dunn, 2011). These considerations of the needs, interests and actions of both the child and family members, illustrates how paediatric rehabilitation intervention can be quite complex.

To advance paediatric rehabilitation intervention research, there is a need to address its unique design challenges since the interventions are often complex and multifaceted (Bartlett et al., 2006). One of the contributing challenges is determining whether key characteristics of paediatric interventions are clearly defined and measured. Once an intervention is defined, a rigorous research trial should measure how accurately the intervention is delivered, a concept described as intervention fidelity (Altman et al., 2001). The purpose of examining intervention fidelity is to ensure that the intervention is implemented as planned (Waltz, Addis, Koerner & Jacobson, 1993). Fidelity monitoring

is typically involved in efficacy studies that aim to evaluate interventions under ideal conditions by controlling the sample selection, intervention delivery and conditions under which the intervention occurred (Hoagwood, Hibbs, Brent & Jensen, 1995).

In the field of paediatric rehabilitation, there is a paucity of rigorous fidelity measures relevant to children with physical disabilities. The objective of this thesis research is to develop and conduct early psychometric testing of a novel fidelity measure in paediatric rehabilitation for children with cerebral palsy. In addition to being relevant to replication research involving interventions for children with cerebral palsy, a fidelity measure can be able to be directly applied in future intervention studies. The measure can also serve as a foundational framework for investigators conducting research on different interventions and/or childhood disability populations.

In the literature, intervention fidelity has many synonyms, including treatment integrity, therapist fidelity, or procedural reliability. Within this thesis, fidelity is defined as intervention fidelity, meaning the degree to which an intervention is implemented as intended (Carroll et al., 2007; Institute of Medicine, 2001; Kendall & Beidas, 2007). Over the last 20 years, the concept of intervention fidelity measurement has evolved as interventions have become more complex. There is currently no model of intervention fidelity that aligns with paediatric rehabilitation, yet the concept of fidelity is discussed in the paediatric literature (Bayona, McDougall, Tucker, Nichols, & Mandich, 2006; Miller, Schoen, James, & Schaaf, 2007; Parham et al., 2011).

The development of fidelity measures is a time and resource consuming process. These measures are necessary to implement in non-pharmacological research but difficult

for busy investigators to develop in addition to mounting an intervention study. To develop a sound measure of fidelity, there is a need to identify all active ingredients of the intervention and a way that objectively measures its complexity. A foundational fidelity measure along with particular methods to adapt it for inclusion in intervention studies would be a useful tool to paediatric rehabilitation researchers. A fidelity measure would also ensure that paediatric rehabilitation moves in directions already being explored by other health disciplines such as in psychology.

This thesis aims to examine the theoretical foundations of fidelity measurement that best align with paediatric rehabilitation, develop a novel measure of fidelity within the context of a specific study and test the reliability and validity of this measure for use by investigators in paediatric rehabilitation.

Defining Intervention in Paediatric Rehabilitation Research

In this era of evidence-based practice, there is an expectation for novel and existing interventions in paediatric rehabilitation to be studied and supported/refuted in part by research evidence. To ensure a high quality of evidence, there is a need for rigorous methods to examine the ways in which intervention trials are implemented. One of the most important factors is a careful and detailed definition of the intervention under examination (Whyte, 2003). Unfortunately, a major limitation of rehabilitation research is the lack of clearly defined interventions (Diijkers et al., 2002). Unlike drug trials, where treatment is precisely prescribed, non-pharmacological paediatric interventions, such as paediatric rehabilitation therapy, are typically dynamic and individualized to meet the needs of the child and their family. The interventions are based on theoretical principles

and tailored to the client and the family within a family-centred practice. The way in which these concepts are operationalized within an intervention session can, thereby, create variation in therapy delivery. It is important to carefully define interventions to ensure that they are delivered as designed, provide predictability of the active ingredients across clinicians and enable the study to be replicable (Whyte & Hart, 2003).

In addition to examining intervention-specific strategies, there has been discussion about other active ingredients of therapeutic intervention that potentially impact outcomes (Abraham & Michie, 2008). Non-specific intervention behaviours commonly employed across therapies may also contribute to changes in human behaviours. These general types of therapy behaviours can take the form of common therapy procedures, such as conducting an assessment. General therapy characteristics also include behaviour change strategies, such as modeling behaviour, setting specific goals, and providing feedback on performance (Abraham & Michie, 2008). Other types of non-specific (or general therapy) behaviours relate to aspects of the therapy process within family-centred practice, whereby therapists' demonstrate help-giving behaviours within an intervention as outlined by Dunst (2009). Examples of these relational characteristics that may impact outcomes include active listening, empathy, and approaches that help the family make decisions (Dunst, 2009). Even though there is no literature in paediatric rehabilitation identifying the impact of these non-specific (or general) therapy behaviours on outcomes, areas in psychology have determined that such strategies (e.g., therapeutic alliance) can influence outcomes (Elvins & Green, 2008). Thus, in the examination of the fidelity of an intervention, the identification of all key

influences within an intervention session on a child's outcome would include these general characteristics of therapy.

The way in which general behaviours of therapeutic interventions impact outcomes provides knowledge about the characteristics that may mediate or moderate intervention. Intervention mediators are considered to be the potential mechanisms or causal links through which an intervention achieves its outcomes, whereas moderators specify conditions influencing the strength or direction of a relationship between the predictor and outcome variables (Kraemer et al., 2002; Rose, Holmbeck, Coakley, and Franks, 2004). The relationship between mediating and moderating factors can be complex and vary depending on the intervention and population. The fundamental descriptions of each concept can be seen in figure 1 (mediating relationship) and figure 2 (moderating relationship). The predictor variable within the figures can be intervention strategies.

[Insert Figure 1 approximately here]

[Insert Figure 2 approximately here]

Based on this understanding of mediating and moderating factors, all characteristics of a therapy intervention can be examined to determine the relationship between the predictor and outcome. Identifying the relationships between specific and non-specific characteristics of therapy with the outcomes will help the understanding of the successful features of the intervention and inform its implementation within clinical practice.

Once the active ingredients of an intervention are defined, manuals are commonly developed to train interventionists within a study to ensure uniformity in therapist intervention delivery, within a study and in future trials (i.e., reproducibility). The manuals describe the intervention's parameters and outline the salient features of the intervention approaches (Chambless & Hollon, 1998; Lambert & Ogles, 2004). The term manual-based intervention is used in the literature to represent these materials. There have been some criticisms that manualizing interventions may cause inflexibility in practitioner delivery of the intervention (Leventhal & Friedman, 2004). However, proponents have emphasized that intervention manuals provide a guideline for clinicians to use the specified intervention within their own therapeutic style rather than suggesting the need to adopt a 'cookbook' approach (Kendall & Beidas, 2007; Lambert & Ogles, 2004).

Fidelity Measurement in Paediatric Rehabilitation

Trends in Health Research

Specific methods of evaluating manual-based interventions, such as measuring a therapist's fidelity to an intervention, are essential to improving the rigor of RCTs in rehabilitation research (Dijkers et al., 2002; Nelson & Mathiowetz, 2004). An expansion of the items in the Consolidated Standards of Reporting Trials (CONSORT) statement to improve the quality of RCTs in non-pharmacologic treatment now includes a section about implementation of intervention as an essential characteristic (Boutron, Moher, Altman, Schultz, & Ravaud, 2008). As a result, evaluation of intervention fidelity is

increasingly important to address across non-pharmacologic health care research in fields such as psychotherapy, nursing, and surgery.

Even though the CONSORT statement extension has emphasized that more intervention details are required in a trial, it has not specifically addressed the problems associated with describing complex interventions. In 2000, the Medical Research Council published a framework (revised in 2008) to better assist researchers to adopt appropriate methods to develop and evaluate complex interventions. The MRC defined complex interventions in health care as comprising:

> "…a number of separate elements which seem essential to the proper functioning of the interventions although the 'active ingredient' of the intervention that is effective is difficult to specify. (…) Complex interventions are built up from a number of components, which may act both independently and interdependently. The components usually include behaviors, parameters of behaviors (e.g., frequency, timing), and methods of organizing and delivering those behaviors (e.g., type(s) of practitioner, setting and location)" (Medical Research Council, 2000).

This framework (Figure 3) provides a systematic way to develop the best intervention and best evaluation methods (Craig et al., 2008). Under "Evaluation" heading, specifically "Understanding Change Processes," this framework explores the ways in which an intervention under investigation is implemented and provides valuable insight as to why an intervention fails or has unexpected consequences or why successful intervention works and how it can be optimized (Craig et al., 2008). It emphasizes the

need to assess fidelity and quality of intervention implementation, as well as help to clarify causal mechanisms and identify contextual factors associated with variation in outcomes.

[Insert Figure 3 approximately here]

To date, the best work in intervention fidelity methodology and measurement has occurred within psychotherapy research. Researchers in psychotherapy have developed and tested methods to address challenges in measuring complex treatment practices and establishing methodological standards for intervention fidelity (Bond, Evans, Salyers, Williams & Kim, 2000). Recently, nursing research has also adopted intervention fidelity methods from psychotherapy to increase the trustworthiness of their interventions (Stein, Sargent & Nicholas, 2007).

In comparison to these disciplines, intervention research in paediatric rehabilitation is equally, if not more, complex and dynamic in its practices since children's communication and motivation to participate differs greatly from adults. In rehabilitation, more specifically occupational therapy, a descriptive paper by Nelson and Mathiowetz (2004), stated that the evaluation of intervention fidelity is an essential characteristic to increase the rigor of occupational therapy (OT) research. In a second descriptive paper, Hennessey & Rumrill (2003) detailed the benefits of evaluating fidelity in rehabilitation intervention. These benefits include improvement in:

- Internal validity (showing a causal link to results from a controlled therapy);
- Construct validity (setting a link between treatment and its theory);

- Statistical power by decreasing type III error, defined as failure to implement a program as planned (Dusenbury et al., 2003); and
- External validity (to replicate methods across research samples/settings).

Measuring Intervention Fidelity in Health Research

Construct of Intervention Fidelity

Fidelity measurement has evolved as investigators articulated key features of their interventions (Baer et al., 2007). The first focus on intervention fidelity emerged out of the early psychotherapy literature by Eysenck (1952). He argued that the vagueness of psychotherapy treatment descriptions in research precluded drawing conclusions about effectiveness and its "…shortcomings highlight the necessity of properly planned and executed experimental studies" (Eysenck, 1952, p.323). By the early 1980's, researchers began to define and develop methods to systematically test intervention fidelity in psychotherapy. As a result, several researchers (Sechrest, West, Philip, Redner, & Yeaton, 1979; Yeaton & Sechrest, 1981; Billingsley, White, & Munson, 1980; and Quay, 1977) made a critical distinction between the treatments intended and the treatment actually delivered, categorized as **treatment integrity** methodology (Lichstein, Riedel & Grieve, 1994). Following this development, research centred on **treatment differentiation** in order to distinguish between treatments (Kazdin, 1980). The two independent concepts of **treatment integrity and differentiation** were studied separately until they were united under the concept of **treatment (or intervention) fidelity** by Moncher & Prinz (1991). Intervention fidelity has been measured in many different ways, ranging from program evaluation of broad procedures of the intervention

(e.g., goal setting, assessment) to measuring session-specific intervention behaviours. Some investigators (Gresham, 1997; Schlosser, 2002) have defined different levels of fidelity reporting as:

- the overall integrity of all intervention components across sessions;
- component integrity of each intervention component across sessions;
- session integrity of all intervention components within an intervention session.

The focus of this thesis is on the development of a measure of intervention session fidelity. Intuitively, there may be a higher degree of difficulty than program evaluation, in implementing and measuring the fidelity of an intervention session in non-pharmacological intervention since the therapist and client receptiveness of the therapy can vary substantially. The types of session specific fidelity measures that are outlined in this section include both intervention-specific and generic fidelity measures.

Intervention-Specific Fidelity Measures

Most instruments measuring intervention fidelity in efficacy research are intervention-specific measures. These measures are developed to examine the active ingredients of a particular intervention across or within intervention sessions (Breitenstein et al., 2010). Within this approach to fidelity measurement, session fidelity has been measured on a macro-level by evaluating the delivery of program components (e.g., goal setting, assessment completed) or on a micro level by measuring therapist implementation of the active ingredients.

The evaluation of program components has been examined in the rehabilitation literature in both mental health and paediatrics. For example, in vocational rehabilitation

(specifically supported employment), a measure of program fidelity was developed by Moll and colleagues (2003) called the Individual Placement and Support model – Fidelity Scale. In paediatric rehabilitation, a measure of program fidelity for school-based occupational therapists was developed to examine therapist ability to complete components of a service protocol (Bayona et al., 2006).

In paediatric rehabilitation, the evaluation of session fidelity of specific intervention behaviours is much less common. Session fidelity of intervention behaviours is challenging to examine, but is important to ensure that the intervention is evaluated beyond whether overall procedures are achieved. Session fidelity assesses the extent and level of quality in how interventions were conducted. This process is crucial in paediatric rehabilitation intervention because the way in which the intervention is implemented is as important as the design of the intervention protocol itself.

Intervention-specific fidelity measures examining session behaviours have been developed in several areas of paediatric rehabilitation. In psychology, fidelity measures have been developed for interventions in autism intervention, such as the Early Start Denver model for young children with autism (Dawson et al., 2010). In occupational therapy, a recent measure was published by Parham and colleagues (2011) to examine intervention fidelity of Sensory Integration Therapy.

Generic fidelity measures

Although intervention-specific measures evaluate therapists' session behaviours, these measures do not examine the characteristics of comparator interventions to ensure its fidelity and to minimize contamination across treatment arms. As well, these measures

do not examine characteristics common to all interventions in a study. Such general therapy behaviours, as stated earlier, can include setting goals, empathy or active listening. To address such issues, psychotherapy researchers have developed 'generic' fidelity measures that include characteristics of more than one intervention and characteristics common to all interventions (Breitenstein et al., 2010). From this definition, it is clear that 'generic fidelity' measures differ in meaning from the traditional meaning of 'generic measures' in instrument development, whereby the items are applicable to all populations and are not specific to any particular interventions (Streiner & Norman, 2006). Generic fidelity measures examine all characteristics of the interventions delivered within a trial to enable investigators to compare the fidelity of the multiple interventions. As a result, each intervention is examined in terms of how well the expected intervention is delivered, and whether any behaviours from the comparator intervention were present. In addition, a few 'generic' fidelity measures in psychotherapy have included therapy characteristics important and common to each intervention in the trial. One example is the Yale Adherence and Competence Scale (Carroll et al., 2000). This measure contains domains for each specific substance abuse intervention involved (e.g., 12-step process, cognitive behavioural therapy and counseling) as well as a domain with non-specific intervention items important to all substance abuse therapies (e.g., goal setting) (Carroll et al., 2000).

Developing a Novel Fidelity Measure in Paediatric Rehabilitation

Intervention-specific fidelity measures place a large burden on individual investigators since they are required to have (a) training manuals (b) training programs

for clinicians in the study, (c) process measures to evaluate therapist fidelity, and (d) preliminary findings of the intervention being studied (Rounsaville & Carroll, 2001). The development of a generic measure with the option to insert intervention-specific items could reduce this burden substantially. A generic measure would enable generalizability across different interventions that are consistent in their conceptual foundations (Breitenstein et al., 2010). Generic measures can reduce the cost of developing new fidelity measures since they provide a foundation or framework to be applied to novel interventions. In **paediatric rehabilitation** research, there is a **scarcity of methods** to measure fidelity in the literature, and there are no **generic fidelity measures.**

Conceptual Models of Intervention Fidelity

This section describes the theoretical frameworks that have formed the basis of measures of intervention fidelity. It emphasizes the importance of identifying characteristics of fidelity that are necessary to evaluate. The section concludes with the identification of the framework most relevant to underpin a novel generic fidelity measure in paediatric rehabilitation.

Traditionally, intervention fidelity has considered fidelity to be the "adherent and competent delivery of an intervention by the interventionist as set forth in the research plans" (Judge-Santacroce, Maccarelli & Grey, 2004, p.63**). Adherence and competence** are typically captured in measuring treatment fidelity within the original Technology Model (to be further discussed in the next section). **Adherence** addresses the extent to which a therapist uses interventions detailed by the treatment manual, while **competence** is the level of skill (or **quality** of skill) demonstrated by the therapist in delivering the

treatment (Waltz et al., 1993).Since the early 1990's, researchers have focused on developing guidelines for fidelity methods and tools to measure components of fidelity relevant to the **unique ingredients of their specific interventions**. As a result, the constructs defining fidelity have advanced over the years. The Technology Model, although still used, has evolved into different models of fidelity including the Treatment Implementation Model and the Framework of Implementation Fidelity.

The Original Technology Model

A prominent fidelity model used in the psychotherapy and nursing literature, the original Technology Model measures only the **therapist's** behaviours (Waskow, 1984). This model assesses therapy delivery in terms of: (1) **dose of treatment** (frequency and number of therapy sessions); (2) **therapy ingredients** (active and inert ingredients of treatment practice); (3) **conditions of treatment** (under which they are administered); and (4) whether the **treatment is delivered to all patients** (adequate therapist compliance). Within each of these areas, fidelity measurement evaluates therapist adherence to an intervention protocol. The Technology model has been cited by 27 studies in the literature (Cited Reference Search #1, 2012). This model is the foundation of the Yale Adherence and Competence Scale (Carroll et al., 2000).

Treatment Implementation Model

A second prominent conceptual model for the measurement of fidelity of intervention is the Treatment Implementation Model (Lichstein et al., 1994), cited 64 times in the literature (Cited Reference Search #2, 2012). The components of this model measure a wider scope of treatment fidelity by considering both the **therapist's**

performance and **client's** behaviours. The three components of the measurement model include: (1) Delivery of treatment; (2) Receipt of treatment; and (3) Enactment of treatment (Lichstein et al., 1994).

Delivery of treatment (or adherence) refers to the therapist's performance in delivering the key therapeutic in the intervention. As well, the model identifies ingredients of associated treatments that are not part of the treatment. This construct does not evaluate competence.

Receipt of treatment assesses how receptive clients are to the therapy in the way in which the therapist implements the treatment. Lichstein postulated that "receipt deficits will always shrink outcome" (Lichstein, et al., 1994, p.12). The therapist and the client have a "shared responsibility" for facilitating success for this component. Receipt is influenced by communication, the characteristics of the therapeutic relationship, motivation, and participation in the treatment (Lichstein et al., 1994).

The third component of this model assesses how well the **treatment is enacted** or practiced by the client outside of a direct therapy session. For example, doing homework to reinforce the therapy provided in the intervention session. Overall, the components of the Treatment Implementation model are defined and measurable but require observations of both the therapist and the client inside and outside of a session.

The Treatment Implementation model has been adopted in recommendations published by the Treatment Fidelity Workgroup of the National Institute of Health Behavior Change Consortium (Bellg et al., 2004). These recommendations describe essential components in the **design and evaluation** of treatment fidelity in research

studies related to behaviour change in health. These essential components include study design, training providers, delivery of treatment, receipt of treatment, and enactment of treatment skills (Bellg et al., 2004).

Framework of Implementation Fidelity

A third model, the Implementation Fidelity framework (IFF) (See Figure 4) includes concepts that integrate many of the characteristics of fidelity about both client receptiveness (referred to as **responsiveness**) and therapist performance (Carroll et al., 2007). This model is based on five key elements from the literature describing implementation fidelity, adherence, exposure or dose, quality of delivery, participant responsiveness, and program differentiation (Dane & Schneider, 1998). The framework of Implementation Fidelity captures constructs beyond the interventionist similar to the Treatment Implementation Model, and integrates specific interventionist concepts such as quality of delivery from the Technology Model. Overall, this model provides a **more comprehensive** framework to evaluate fidelity at a session level since it focuses on how treatment is conducted and received, and **the relationship** between these constructs. Therapist **adherence mediates** between an intervention plan and fidelity evaluation, and is distinct from the **potential moderators** of intervention such as client responsiveness and quality of therapy delivery. The framework recognizes that all of these factors can influence fidelity of intervention. The IFF has been cited 19 times in the literature, but has not yet been used to underpin the development of a measure (Cited Reference Search #3, 2012).

[Insert Figure 4 approximately here]

Framework that best fits needs for paediatric rehabilitation

The selection of a particular framework for the development of a fidelity measure should be based on the fit between characteristics of the intervention itself and the constructs of the framework. Each of the aforementioned fidelity models could be used as the conceptual foundation of fidelity measures in paediatric rehabilitation. However, in principle, the Implementation Fidelity Framework provides the most comprehensive constructs to guide the evaluation of **fidelity in a paediatric rehabilitation session**. Within this framework, the key factors potentially influencing fidelity include the dynamic nature of intervention involving the child and/or parent (i.e., client responsiveness) and the therapist's performance (i.e., adherence and quality of intervention). Previous research indicates the importance of including a component of patient responsivity to account for how well an intervention is implemented (Morgenstern & McKay, 2007).

Client responsiveness: Within an intervention session children can have an impact on therapeutic outcomes through their **responsiveness** regardless of how well the therapist adhered to the intervention. Responsiveness to intervention is essential for the intervention to be effective, given that low motivation and inattentiveness can minimize the benefits of intervention (Nelson & Mathiowetz, 2004). Paediatric rehabilitation is a dynamic process in which the client plays a participatory role (Bartlett et al., 2006). Responsiveness in the IFF model indicates that it is not only a moderator influenced by the individuals receiving the intervention, but also influenced by those responsible for delivering intervention (Carroll et al., 2007). The example provided by Carroll et al.,

2007(p.16) was that "higher levels of implementation fidelity were achieved when those responsible for delivering an intervention were enthusiastic about it". This construct is appropriately suited for paediatric intervention, since the therapist can have an influence on the motivation and attentiveness of the child (among other attributes).

Therapist Performance: Another important consideration in paediatric intervention is the therapist's performance in an intervention session. Areas of therapist performance considered essential for fidelity are **adherence** and **quality of intervention delivery (competence)**. Adherence includes both adhering to the **active ingredients** of a specific intervention, as well as not implementing ingredients that align with other interventions (to **differentiate** each approach). Quality of intervention focuses on intervention being conducted in a skillful way and in an appropriate and timely manner (Carroll et al., 2000).

Each of these essential factors for both child and therapist characteristics are captured comprehensively within the Implementation Fidelity framework. In Table 1, the frameworks are compared to assess overall fit for paediatric rehabilitation intervention. Each model shares some complementary principles, however each is presented with its own unique language. From this analysis, the Implementation Fidelity framework is the most comprehensive fit for the important attributes of paediatric rehabilitation.

[Insert Table 1 approximately here]

Although the **Implementation Fidelity framework** labels other moderators that impact fidelity, the characteristics of adherence to intervention, quality of delivery and participant responsiveness are the attributes that can be evaluated in an intervention

session. Moderators such as 'comprehensiveness of policy description', and 'strategies to facilitate implementation' are more important for study design than session evaluation. A fidelity measure that assesses the responsiveness of the child and the therapist's adherence and quality of intervention will provide a comprehensive evaluation of session intervention fidelity.

In using the constructs **of implementation fidelity** in this project, therapy **adherence** can mediate intervention fidelity, but can be strengthened (or weakened) by the **quality of therapy delivery** and **child responsiveness**. The **definition of adherence** in this model focuses on evaluating the active ingredients of the intervention. However, **differentiation** between competing interventions in an RCT is also important to evaluate given the foundational understanding of fidelity developed by pioneers Moncher & Prinz (1991) and Dane & Schneider (1998). High levels of fidelity are achieved when:

- Therapist adherence includes the active ingredients of the intervention without being contaminated by ingredients that are not part of intervention;
- Intervention is conducted with high quality of delivery; and
- Child/Parent-therapist interaction demonstrates a high level of client responsiveness.

In summary, there is a need for rigorous evidence-based paediatric interventions that includes thoughtful attention in the evaluation of intervention fidelity. The measurement of fidelity in paediatric rehabilitation so far is not extensive. There is no fidelity measure developed for inter-professional paediatric rehabilitation. None of the existing fidelity measures in the paediatric literature have considered both the evaluation

of the therapist and client responsiveness during intervention. Thus, research to develop reliable and valid measurement of fidelity in paediatric rehabilitation is needed.

Research Objective

The objective of this thesis is to identify observable and important general characteristics of paediatric rehabilitation to develop and test a novel generic fidelity measure. The measure will be developed within the context of an existing intervention study.

Context for the thesis

The development of the fidelity measure was completed using data from a study titled, 'Family centred functional therapy for children with cerebral palsy' (referred to as the Focus on Function Study) (Law et al., 2011). The primary aim of the Focus on Function (FOF) study was to conduct a multi-site RCT to evaluate the efficacy of two approaches to paediatric rehabilitation intervention. These interventions aimed to improve performance of functional tasks and mobility and increase participation in everyday activities in 128 children with cerebral palsy (12 months to 5 years of age). Within the study, therapists were randomly allocated to provide interventions using one of two manualized rehabilitation approaches to improve functional performance. The children within the study followed their randomized therapist into one of the following interventions: (i) Child-focused approach, i.e., changing the child's skills and abilities; and (ii) Context-focused approach, i.e., changing the task or environment supporting performance.

Within the Focus on Function study, over 70 occupational therapists (OT) and physiotherapists (PT) across 21 children's centres in Ontario and Alberta, Canada conducted the two interventions. In both intervention arms, children were seen by an OT or PT as their main therapist within the trial. Other therapy disciplines were typically available for consultation (Context-focused approach) or actively worked with the child (Child-focused approach) depending on the child's intervention goals. Prior to conducting the interventions, all therapists underwent face-to-face training for their assigned intervention and had a manual developed by Drs. Law and Darrah (Co-PI's) and their research team. The materials from the study used in this thesis included videotaped therapy sessions (one session per child), and therapist session clinical chart notes (across all intervention sessions).

Thesis Goals

This introductory chapter provides the background and purpose of a thesis that describes the development of a novel fidelity measure for paediatric rehabilitation. The chapters within this thesis each provide new knowledge in addition to building on each other to develop the generic fidelity measure titled the Paediatric Rehabilitation Observational measure of Fidelity (PROF). Chapter two describes a narrative review of the literature that identifies existing generic fidelity measures and the relationship of therapy process items in these measures to paediatric rehabilitation. Chapter three presents the findings from a Delphi Process conducted with experts in paediatric rehabilitation to identify essential characteristics of the therapeutic process in paediatric rehabilitation intervention for children with physical disabilities. Chapter four describes

the development and initial psychometric testing of the new fidelity measure, the PROF. Chapter five focuses on a discussion of all findings and the overall conclusion of the thesis, implications of these findings and ideas for future research.

Ethical Considerations

To conduct chapters 3 and 4 for this thesis, ethical approval was necessary and was secured from the Hamilton Health Science Faculty of Health Sciences Research Ethics Board at McMaster University.

In chapter 3, two processes were conducted to generate data – a Delphi process and semi-structured interviews. The Delphi process involved eight multidisciplinary health care researchers and rehabilitation clinicians that participated in an online consensus process to identify essential attributes of general therapy behaviours in paediatric rehabilitation. In addition to voting and providing their perspective for each of the three iterations of the list of attributes, the participants shared professional demographic information. Informed consent was collected and procedures within the online exercise were implemented to secure the anonymity of the eight participants.

For the semi-structured interviews, seventeen practicing occupational therapists and physiotherapists from the Focus on Function (FOF) study were recruited to (through the FOF research coordinator) to participate in an interview. The interviews consisted of a one-time tape-recorded telephone interview with BD. Participants were assured that they did not have to answer any questions that they did not want to answer. These interviews were transcribed and participant names were removed from the transcripts and

the results of each interview were summarized. Informed consent was collected and anonymity was secured for participants.

In chapter 4, the study to develop and test the psychometrics of the PROF, two sources of data required ethical considerations – demographic information for the video data and independent raters. The video data used to test the properties of the PROF were a secondary data source from the FOF study. The children, parents and therapists who were on video in this secondary data source gave informed consent to use the material for secondary analysis within the FOF study. These data consisted of intervention sessions involving therapists and the child (and sometime the parent) participating in the intervention. All personal identifying information related to the child and therapist demographic information was removed by BD and assigned a numeric identifier to protect the identity of all participants involved.

The psychometric testing of the PROF involved the use of six independent raters who were practicing occupational therapists or physiotherapists in paediatric rehabilitation in southern Ontario. Informed consent was collected and anonymity was secured for all raters. In addition, participants committed to maintaining the confidentiality of the identity of therapists and/or children on video that they may be familiar with, since the FOF study involved clinical sites in Ontario.

References

Abraham, C., & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. Health Psychology, 27, 379-387.

Allen, K.D., Wilczynski, S. M., and Evans, J.H. (1997). Pediatric rehabilitation: Defining a field, a focus and a future. *International Journal of Rehabilitation and Health, 3,* 25-40.

Altman, D.G., Schulz, K,F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P.C., Lang, T.; CONSORT GROUP (Consolidated Standards of Reporting Trials). (2001). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine, 134,* 663-94.

Baer, J.S., Ball, S.A., Campbell, B.K., Miele, G.M., Schoener, E.P., and Tracy, K. (2007). Training and fidelity monitoring of behavioral interventions in multi-site addictions research. *Drug and Alcohol Dependence, 87*, 107-118.

Bartlett, D.J., Macnab, J., Macarthur, C., Mandich, A., Magill-Evans, J., Young, N.L., Beal, D., Conti-Becker, A., & Polatajko, H. (2006). Advancing rehabilitation research: an interactionist perspective to guide question and design. *Disability and Rehabilitation, 28,* 1169-76.

Bayona, C.L., McDougall, J., Tucker, M.A., Nichols, M., & Mandich, A. (2006). School-based occupational therapy for children with fine motor difficulties: Evaluating functional outcomes and fidelity of services. *Physical & Occupational Therapy in Pediatrics, 26,* 89-110.

Bellg, A.J., Borrelli, B., Resnick, B., Hecht, J., Sharp Minicucci, D., Ory, M. et al. (2004). Enhancing treatment fidelity in health behaviour change studies: Best practices and recommendations from the NIH Behaviour Change Consortium. *Health Psychology, 23,* 443-451.

Billingsley, F., White, O.R., & Munson, R. (1980). Procedural reliability: A rationale and an example. *Behavioral Assessment,* 2, 229-241.

Bond, G. R., Evans, L., Salyers, M., Williams, J., & Hea-Won, K. (2000). Measurement of fidelity in psychiatric rehabilitation. *Mental Health Services Research, 2,* 75-87.

Boutron, I., Moher, D., Altman, D.G., Schultz, K.F., & Ravaud, P. (2008). Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment: Explanation and Elaboration. *Annals of Internal Medicine, 148,* 295-306.

Breitenstein, S.M., Gross, D., Garvey, C.A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation Fidelity in Community-Based Interventions. *Research in Nursing & Health, 33,* 164-173.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science 2*, 40 doi:10.1186/1748-5908-2-40

Carroll, K.M., Nich, C., Sifry, R.L., Nuro, K.F., Frankforter, T.L., Ball, S.A. (2000). A general system for evaluating therapist adherence and competence in psycho-therapy research in the addictions. *Drug and Alcohol Dependence, 57,* 225-238.

Chambless, D.L., & Hollon, S.D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66,* 7-18.

Cited Reference Search #1: Cited Reference Index for Waskow, IE (1984) from Psychotherapy Research. Retrieved on April 16, 2012, from the Web of Science database.

Cited Reference Search #2: Cited Reference Index for Lichstein KL et al. (1994) from Advances in Behaviour Research and Therapy. Retrieved on April 16, 2012, from the Web of Science database.

Cited Reference Search #3: Cited Reference Index for Carroll C et al. (2007) from Implementation Science. Retrieved on April 16, 2012, from the Web of Science database.

Craig, P., Dieppe, P., Macintyre, S., Mitchie, S., Nazareth, S., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal, 337,* 979-983.

Dane, A., & Schneider, B. H. (1998). Integrity in primary prevention programs: Are implementation effects out of control? *Clinical Psychology Review, 18,* 23-45.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., Donaldson, A., & Varley, J. (2010). Randomized, Controlled Trial of an Intervention for Toddlers With Autism: The Early Start Denver Model. *Pediatrics, 125,* e17-e25.

Dijkers, M., Kroop, G.C., Esper, R.M., Yavuzer, G., Cullen, N., Bakdalieh, Y. (2002). Quality of intervention research reporting in medical rehabilitation journals. *Amer J Phys Med, 81,* 21-33.

Dunn, W. (2011). Best practice occupational therapy for children and families in community settings. Thorofare, NJ: Slack

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. (2003) A Review of Research on Fidelity of Implementation: Implications for Drug Abuse Prevention in School Settings. *Health Education Research 18,* 237-256.

Dunst, C.J. (2009). Implications of Evidence-Based Practices for Personnel Preparation Development in Early Childhood Intervention. *Infants and Young Children, 22,* 44-53.

Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28,* 1167-1187.

Eysenck, H. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16*, 319-324.

Gresham, F.M. (1997). Treatment integrity in single-subject research. In R. D. Franklin, D.B. Allison, & B.S. Gorman (Eds.), *Design and analysis of single-case research* (pp.93-117). Mahwah, NJ: Erlbaum.

Hennessey, M.L., & Rumrill, P.D. (2003). Treatment fidelity in rehabilitation research. *Journal of Vocational Rehabilitation, 19,* 123-126.

Hoagwood, K., Hibbs, E., Brent, D., and Jensen, P. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 683-687.

Institute of Medicine (2001). *Improving the quality of long-term care.* National Academy Press, Washington, DC, USA.

Judge Santacroce, S.J., Maccarelli, L.M. & Grey, M. (2004). Intervention fidelity. *Nursing Research, 53,* 63-66.

Kazdin, A.E. (1980). *Research design in clinical psychology.* New York: Harper & Row.

Kendall, P.C., & Beidas, R.S. (2007). Smoothing the trail for dissemination of evidence-based practices for youth: Flexibility within fidelity. *Professional Psychology: Research and Practice, 38,* 13-20.

Kraemer, H., Wilson, T., Fairburn, C., & Agras, W.S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59,* 877-883.

Lambert, M. J., &, Ogles, B. M. (2004). The Efficacy and Effectiveness of Psychotherapy. In M.J. Lambert, *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (5th ed.) pp. 139–193. New York: John Wiley & Sons.

Law, M., Darrah, J., Pollock, N., Wilson, B., Russell, D., Walter, S., Rosenbaum, P., Galuppi, B. (2011). Focus on function: a cluster, randomized controlled trial comparing child- versus context-focused intervention for young children with cerebral palsy. *Developmental Medicine and Child Neurology,* 53, 621-629.

Leventhal, H.E., & Friedman, M.A. (2004). Does Establishing Fidelity of Treatment Help Us to Understand Treatment Efficacy? *Health Psychology, 23,* 452-456.

Lichstein, K.L., Riedel, B., & Grieve, R. (1994). Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research & Therapy, 16,* 1-29.

Medical Research Council (2000). A framework for development and evaluation of RCTs for complex interventions to improve health 2000. [ONLINE] Available at http://www.mrc.ac.uk/complex-interventionsguicance. Accessed on April 16, 2012.

Miller, L. J., Schoen, S. A., James, K., & Schaaf, R. C. (2007). Lessons learned: A pilot study on occupational therapy effectiveness for children with sensory modulation disorder. *American Journal of Occupational Therapy, 61,* 161–169.

Moll, S., Huff, J., & Detwiler, L. (2003). Supported employment: Evidence for a best practice model in psychosocial rehabilitation, *70*, 298-310.

Moncher, F.J., & Prinz, R.J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11,* 247-266.

Morgenstern, J., & McKay, J.R. (2007). Rethinking the paradigms that inform behavioral treatment research for substance use disorders. *Addiction, 102,* 1377-1389.

Nelson, D.L., & Mathiowetz, V. (2004). Randomized controlled trials to Investigate occupational therapy research questions. *American Journal of Occupational Therapy, 58,* 24-34.

Parham, D., Roley, S., May-Benson, T., Koomar, J., Brett-Green, B., Burke, J.P., Cohn, E., Mailloux, Z., Miller, L.J., Schaaf, R., (2011). Development of a fidelity measure for research on the effectiveness of the Ayres Sensory Integration intervention. *American Journal of Occupational Therapy, 65,* 133-142

Quay, H.C. (1977). The three faces of evaluation: What can be expected to work. *Criminal Justice and Behavior, 4,* 341-354.

Rodger, S., & Keen, D. (2010). Child- and Family-Centred Service Provision. In S. Rodger (Ed.), *Occupation-centred Practice with children* (pp. 45-69). West Sussex, UK: John Wiley & Sons, Inc.

Rose, B., Holmbeck, G., Coakley, R., & Franks, E. (2004). Mediator and moderator effects in developmental and behavioral pediatric research. *Developmental and Behavioral Pediatrics, 25,* 58-67.

Rounsaville, B.J., & Carroll, K.M. (2001). A stage model of behavioral therapies research: Getting started and moving on from stage 1. *Clinical Psychol Sci Prac, 8,* 133-142.

Schlosser, R.W. (2002). On the importance of being earnest about treatment integrity. *Augmentative and Alternative Communication, 18,* 36-44.

Sechrest, L., West, S.G., Philips, M.A., Redner, R., & Yeaton, W. (1979). Some neglected problems in evaluation research: Strength and integrity of treatments. In L. Sechrest, S.G. West, M.A. Philips, R. Redner, & W. Yeaton (Eds.), *Evaluation studies review annual* (Vol.4, pp. 15-35). Beverly Hills: Sage.

Stein, K.F., Sargent, J.T., & Nicholas, R. (2007). Intervention research: Establishing fidelity of the independent variable in nursing clinical trials. *Nursing Research, 56,* 54-62.

Streiner, D.L. & Norman, G.R. (2006). *Health Measurement Scales: A practical guide to their development and use 3rd Ed.* New York: Oxford University Press.

Waltz, J., Addis, M.E., Koerner, K. & Jacobson, N.S. (1993). Testing the integrity of psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61,* 620-630.

Waskow, I.E. (1984). Specification of the technique variable in the NIMH Treatment of Depression Collaborative Research Program. In: Williams, J.B.W., Spitzer, R.L. (Eds.), Psychotherapy Research: Where are we, and where should we go? New York: Guilford.

Whyte, J. (2003). Clinical trials in rehabilitation: What are the obstacles? *American Journal of Physical Medicine and Rehabilitation, 82,* 10(Suppl), S16-S21.

Whyte, J. & Hart, T. (2003). It's more than a black box; It's a Russian doll: Defining rehabilitation treatments. *American Journal of Physical Medicine and Rehabilitation*, 82, 639-652.

Yeaton, W.H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity and effectiveness. *Journal of Consulting and Clinical Psychology, 49,* 156-167.

Table 1. Summary of key attributes to evaluate fidelity of Paediatric Rehabilitation in relation to prominent intervention fidelity models

	Relevant Concepts from intervention fidelity models		
Fidelity Attributes to evaluate Paediatric Rehabilitation	**Technology Model**	**Treatment Implementation**	**Implementation Fidelity Framework**
Therapist adherence to intervention principles	• Dose • Ingredients • Conditions • Delivery to all patients	Delivery of treatment	• Content • Coverage (i.e., dose) • Frequency • Duration
Quality	Competence (i.e., Skillfulness)	Not addressed	Appropriate Delivery as intended **(related to skillfulness rating)**
General Paediatric Practice Competencies	Not addressed	Receipt of Treatment	Participant Responsiveness

Figure 1. Mediated Relationship

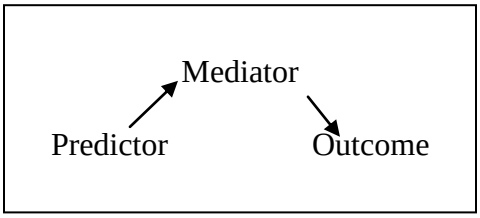


Figure 2. Moderated Relationship

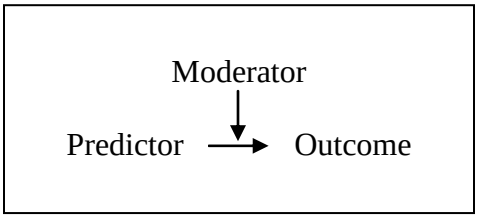


Figure 3. **Updated Medical Research Council Framework**

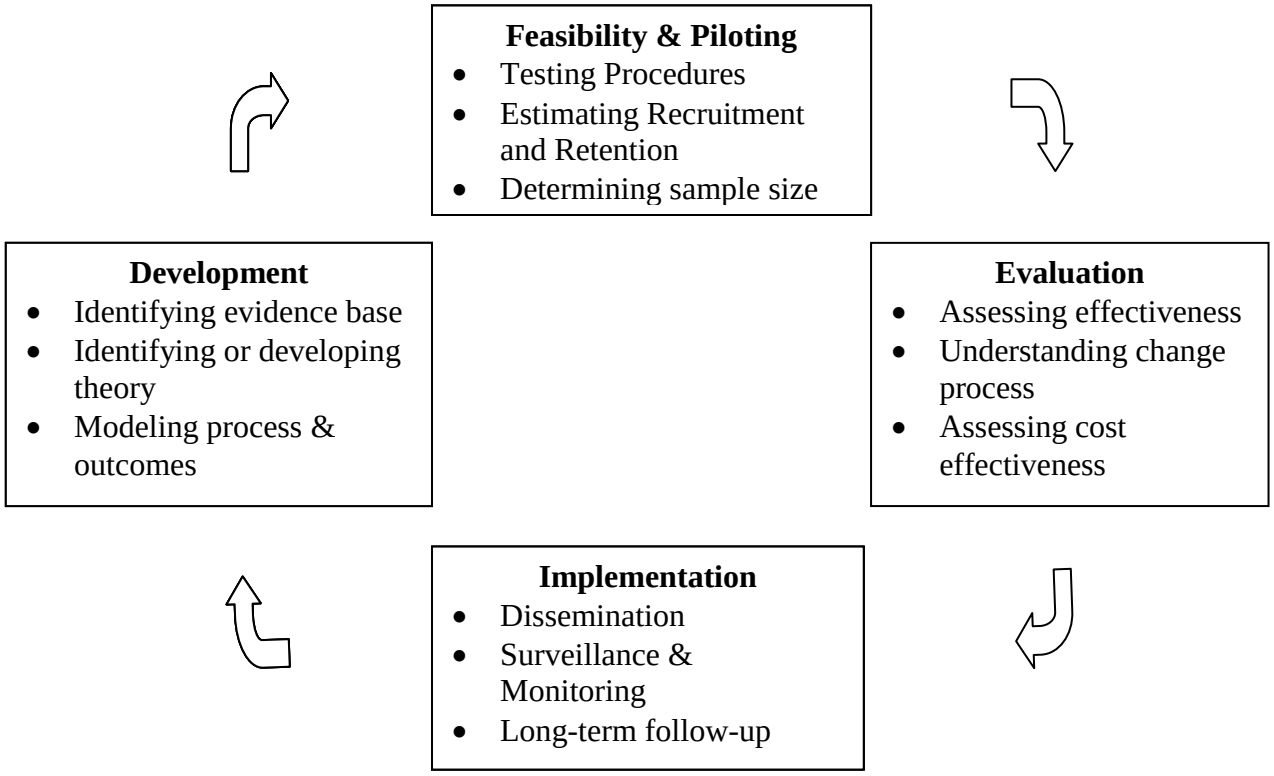


Figure 3. Adapted from "Developing and evaluating complex interventions: the new Medical Research Council guidance'" by P. Craig, P. Dieppe, S. Macintyre, S. Mitchie, I. Nazareth, and M. Petticrew, 2008, *British Medical Journal, 337,* p.980. Copyright 2008 by the BMJ Publishing Group Ltd.

Figure 4. Conceptual framework for implementation fidelity

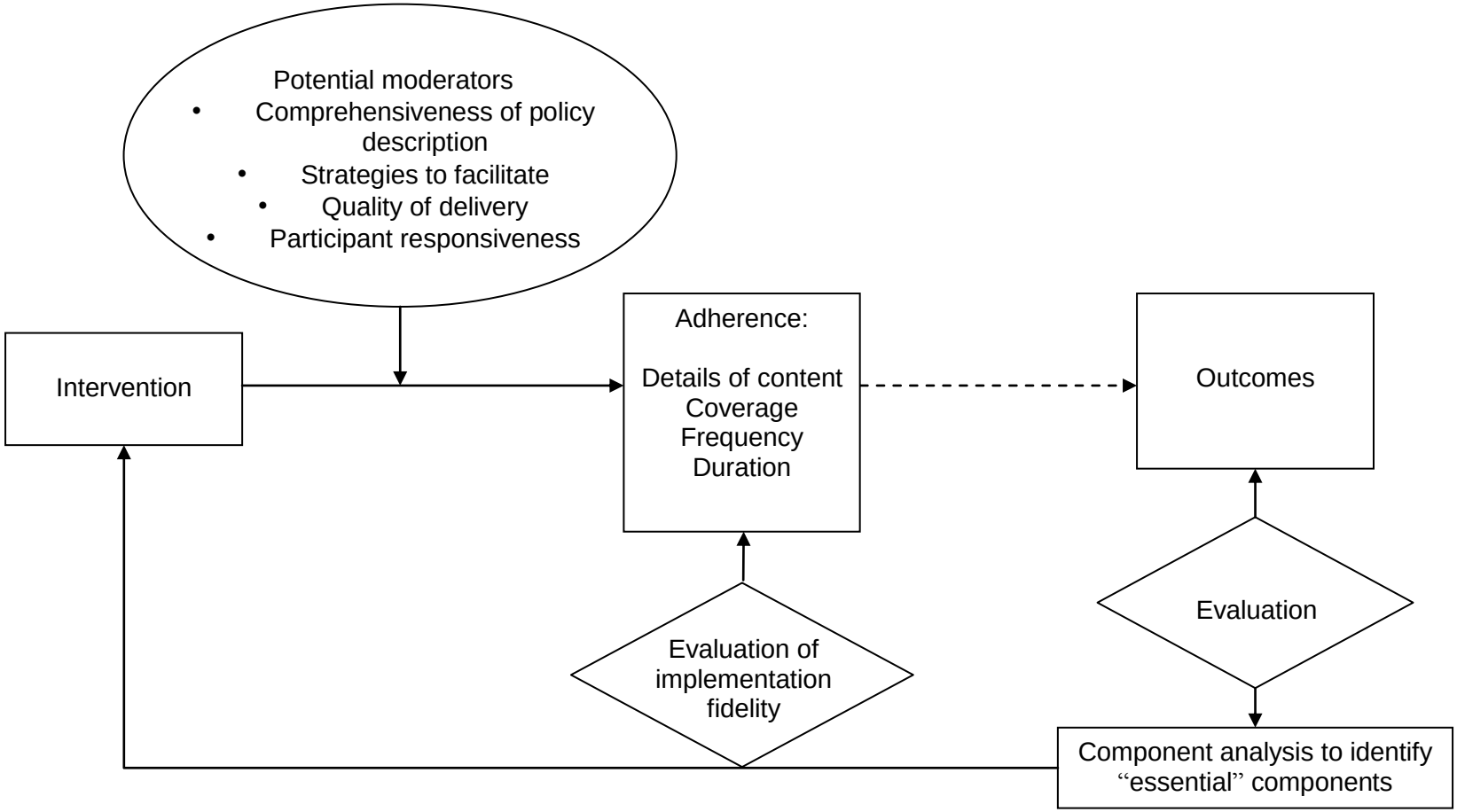


Figure 4. Adapted from "A conceptual framework for implementation fidelity," by C. Carroll, M. Patterson, S. Wood, A. Booth, J. Rick, and S. Balain, 2007, *Implementation Science, 2:* 40, p.4, doi:10.1186/1748-5908-2-40. Copyright 2007 by BioMed Central Ltd.

Chapter Two

A Narrative Review of Generic Intervention Fidelity Measures

Authors: Briano Di Rezze, Mary Law, Jan Willem Gorter, Kevin Eva, Nancy Pollock

This chapter contains a manuscript entitled "A narrative review of generic intervention fidelity measures". This work is foundational within the field of intervention fidelity measurement by highlighting the characteristics of generic fidelity measures. Nowhere in the health and education literature, is there a review of the characteristics of current generic fidelity measures. This review is also essential in providing the foundation of the subsequent studies in this thesis that amount to the generation of the first generic fidelity measure in paediatric rehabilitation.

This manuscript has been accepted for publication in the Journal of Physical and Occupational Therapy in Pediatrics on May 9, 2012. Written confirmation for the acceptance of this paper is below, as well as permission to reproduce this manuscript within this thesis.

Subject: Physical & Occupational Therapy In Pediatrics - Decision on Manuscript ID WPOP-2011-0076.R2
Date: Wed, 9 May 2012 18:52:06 -0400 (EDT)
To: direzzbm@mcmaster.ca

09-May-2012

Dear Mr Di Rezze:

Ref: A narrative review of generic intervention fidelity measures

We are pleased to accept your paper in its current form which will now be forwarded to the publisher for copy editing and typesetting. The reviewer comments are included at the bottom of this letter.

You will receive proofs for checking, and instructions for transfer of copyright in due course.

The publisher also requests that proofs are checked and returned within 48 hours of receipt.

Thank you for your contribution to Physical & Occupational Therapy In Pediatrics and we look forward to receiving further submissions from you.

From: "Palisano,Robert" <rjp33@drexel.edu>
Subject: POTP manuscript
Date: Wed, 9 May 2012 19:56:34 -0400
To: "direzzbm@mcmaster.ca" <direzzbm@mcmaster.ca>
Cc: "Palisano,Robert" <rjp33@drexel.edu>, "Mitchell,Richard" <rbm38@drexel.edu>

Briano:

You have permission to include your manuscript *A Narrative Review of Generic Intervention Fidelity Measures*

In your dissertation.

Robert J. Palisano, PT, ScD, FAPTA

Co-Editor POTP

Drexel University

Physical Therapy and Rehabilitation Sciences

245 N. 15th Street

Philadelphia, PA 19102-1192

Phone: 215-762-1006

Fax: 215-762-3886

E-mail: robert.j.palisano@drexel.edu

Abstract

To increase the rigour of paediatric rehabilitation research, there is a need to evaluate the degree to which an intervention is conducted as planned (i.e., fidelity). Generic fidelity measures evaluate more than one intervention and often include non-specific attributes of the therapy process common to both interventions. The objective of this study was to describe the characteristics of generic fidelity measures and examine how these attributes fit with paediatric rehabilitation. A review of generic fidelity measures was conducted utilizing health and education databases. Five generic fidelity measures are described and examined for their applicability to paediatric rehabilitation. The measures were used in nine studies meeting the inclusion criteria, involving people ages 11 to > 65 years undergoing psychotherapy. From the 76 non-specific items, 37 items were judged to be applicable to paediatric rehabilitation. Common characteristics of non-specific attributes with paediatric rehabilitation are discussed, and investigator plans to conduct future testing.

Keywords: Fidelity, fidelity measures, procedural reliability, intervention studies, health services, rehabilitation, physical therapy [intervention], occupational therapy [intervention].

Introduction

Intervention fidelity is defined as the degree to which a therapist implements an intervention as intended (Carroll et al., 2007; Kendall & Beidas, 2007). The concept of intervention fidelity originated in the field of psychotherapy and gained momentum when the term "treatment fidelity" was first coined by Moncher & Prinz (1991). Since then, methods of examining fidelity in treatment (or intervention) have been used to strengthen the validity of research on complex interventions through documenting a causal link between a defined therapy and outcomes. Evidence of fidelity ensures a close relationship between intervention and its theory and replication of methods across research samples/settings (Hennessey & Rumrill, 2003). Incorporating methods of defining intervention and evaluating fidelity in paediatric rehabilitation research is important to examine whether an intervention is administered as planned (Whyte, 2003). In addition, fidelity measures may have applications that benefit practitioners and students to learn and/or evaluate their delivery of specific interventions.

Traditionally, the examination of intervention fidelity has focused on evaluating therapist adherence to the treatment as a mediating factor between the intervention and outcomes (Waltz, Addis, Koerner & Jacobson, 1993). The construct of fidelity has evolved to articulate key features of unique and increasingly complex interventions emerging in health care practice (Baer et al., 2007). In addition to defining the unique features of an intervention, other concepts have become increasingly important to consider. These concepts include differentiation between comparator interventions and the identification of non-specific attributes that are essential but common to all

interventions (Calsyn, 2000). As a result, models describing fidelity constructs have increased in complexity.

A recent model of fidelity, the Implementation Fidelity Framework (IFF), outlines potential factors that both mediate and moderate intervention fidelity (Carroll et al., 2007). Factors that mediate intervention fidelity focus on therapist adherence to delivering the intervention as expected and include only therapist behaviours. Moderating factors of intervention fidelity examine concepts beyond the therapist's adherence to the intervention. Examples include therapist quality of intervention delivery, client attributes independent of the therapist, and client attributes dependent on the therapist (i.e., therapist-client interaction). Although the IFF is not the first model to examine moderators of fidelity (Waltz et al., 1993), this model includes moderators relevant to the client behaviours. Other literature has emphasized the importance of client responsiveness as a factor that can influence the degree of fidelity with which an intervention is implemented (Morgenstern & McKay, 2007).

Paediatric rehabilitation can be complex to study because there are many factors influencing therapist performance in an intervention. The involvement of children and families within an intervention are important factors that could influence therapist adherence. In addition to being child-centred, key service provider behaviours are detailed in a conceptual framework of Family-Centred Services (FCS) (Rosenbaum, King, Law, King & Evans, 1998). The FCS is based on the concept of client-centred practice (or person-centred therapy) developed by psychologist Carl Rogers. Behaviours described by the FCS include assisting the family to identify strengths, collaborating with

parents, communicating clearly, supporting families, and considering the psychosocial needs and encouraging the participation of all family members.

The inclusion of factors that moderate fidelity in the IFF aligns well with the dynamic nature of paediatric rehabilitation intervention provided by occupational therapists and physiotherapists. The responsiveness of younger children (0-11 years) and/or the parent within an intervention session will likely have an impact on how and what the therapist delivers in terms of intervention. Since part of what paediatric rehabilitation therapists do in an intervention session is client driven, a measure of intervention fidelity in paediatric rehabilitation will need to consider both mediating and moderating factors.

Rigorous measures of intervention fidelity have been developed by investigators in other disciplines. Two types of measures of intervention fidelity are evident in the literature – *specific* and *generic* measures. Specific measures examine therapist adherence to one intervention or the actual delivery of the program. Generic measures evaluate both the adherence and differentiation of more than one intervention in an area of practice (e.g., psychotherapy) or client population (e.g., substance users) (Breitenstein et al., 2010). In addition, some generic fidelity measures have also included non-specific characteristics common to all interventions, such as therapy process (e.g., assessment) and delivery process (e.g., rapport).

Both specific and generic measures are useful approaches to evaluating intervention fidelity and the appropriate type to use depends on the research question. Specific measures, since they only measure behaviours for one intervention, are best

suited for a study examining the active ingredients of a novel intervention. Generic measures are used to evaluate the fidelity of more than one intervention, where the attributes for both interventions are clearly defined.

To date, there have only been two specific fidelity measures published in the paediatric rehabilitation literature (Bayona, McDougall, Tucker, Nichols & Mandich, 2006; Parham et al., 2011). They are used to demonstrate program evaluation (Bayona et al., 2006) and delivery of intervention (Parham et al., 2011) for a single intervention. These measures do not include moderators influencing the degree of fidelity (i.e., child responsiveness, or non-specific intervention skills) or attributes that differentiate two or more interventions from each other. Thus, a review of measures from other clinical areas may inform the development of a generic fidelity measure in paediatric rehabilitation. The purpose of this paper is to report findings from a review of the characteristics of generic fidelity measures from other clinical areas that could be useful in the development of a measure for paediatric rehabilitation.

Methods

Narrative reviews examine a broad range of issues that describe cutting-edge developments in an area of practice or research (Cook, Mulrow, & Haynes, 1997). As outlined by Roberts (2007), the steps used to conduct the narrative review are to (1) systematically extract the necessary studies; (2) check the extracted studies to ensure the search is complete; and (3) conduct a narrative synthesis of the study content.

Search Strategy

A detailed list of key words was developed through an iterative process of searching the literature with a range of synonyms used for intervention fidelity. Synonyms of the phrase "Intervention Fidelity" included: "Fidelity" AND "Treatment OR Intervention OR Implementation OR Therapy OR Clinician OR Therapist OR Interventionist"; "Integrity" AND "Intervention OR Treatment OR Program"; "Adherence OR Differentiation OR Compliance" AND "Clinician OR Therapist"; "Procedural Reliability"; and "Independent Variable Accuracy." The databases used to search the literature (years 1980-2011) were: CINAHL; EMBASE; Health and Psychosocial Instruments (HAPI); MEDLINE; PsychInfo; and ERIC.

Due to the different order in which the term fidelity is used across the literature (e.g., intervention fidelity or fidelity of intervention), the planned keyword search strategy, with the help of a health sciences librarian (NB, see Acknowledgements), utilized command functions unique to each database to capture the different orders the terms can appear, as well as the range of proximity that may occur between the search terms (e.g., treatment fidelity or fidelity of therapists performing treatment).

The studies were screened using the following inclusion and exclusion criteria. The inclusion criteria consisted of: studies about the development of a measure of intervention fidelity; fidelity measures that differentiate more than one intervention; interventions that are health care related; and measures that also include the examination of non-specific intervention characteristics. Studies were excluded that: focused on one intervention only (i.e., intervention-specific measures); were a thesis; were unable to be

accessed; or the reported trial results described fidelity procedures with insufficient detail.

To ensure that all appropriate literature was captured, a secondary citation search was conducted by searching the literature for references of conceptual papers on intervention fidelity. Searches were conducted using the "citation search" procedure within the Web of Science database. Papers used for this search included: Waskow (1984); Moncher and Prinz (1991); Waltz et al. (1993); Lichstein et al. (1994); Dane and Schneider (1998); and Carroll et al. (2007).

Narrative Synthesis

From the selected studies, descriptive characteristics of the measures were extracted and summarized and psychometric properties detailed**.** A narrative critique was conducted to examine the constructs, domains and scales of the fidelity measures for mediating and/or moderating factors. A second narrative critique examined the non-specific items of the measures and their potential applicability to paediatric rehabilitation intervention. The criteria used to judge item applicability to paediatric rehabilitation was based on the Family-Centred Service framework (Rosenbaum et al., 1998). An agreement procedure was conducted between two investigators (BD and ML) to ensure non-specific items are potentially applicable to paediatric rehabilitation. Each investigator independently reviewed the list of items and recorded whether it was applicable to paediatric rehabilitation. Upon completion, the investigators shared their responses to determine agreement, expected to be greater than 80%. Consensus discussion was used to make decisions about unresolved items for the final list of attributes.

Results

Descriptive results

A total of nine studies were selected. Since some studies used different versions of the same measure, five distinct generic fidelity measures were identified (see Figure 1). The secondary citation search yielded no new studies. The types of studies that were excluded were primarily studies that involved intervention-specific measures; several studies reported generic measures that did not include non-specific items.

[Insert Figure 1 here]

Seven of the studies identified different versions of three generic fidelity measures. For this review, these measures were accessed from the original authors. The *Therapist Behavior Rating Scale – 4th Version* (TBRS-4) by Hogue and Samuolis (2000) was used to describe an intervention for the family targeting drug use and related behavioural problems of adolescents (Hogue, Liddle, Singer & Leckrone, 2005; Diamond, Diamond & Hogue, 2007). The *Collaborative Study Psychotherapy Rating Scale - 6th Version* (CSPRS-6) by Hollon (1984) was used to describe psychotherapy intervention using cognitive-behavioural and interpersonal therapy (McIntosh et al., 2005; Markowitz et al., 2000; Hill, O'Grady & Elkin, 1992). The *Yale Adherence and Competence Scale - 2nd Ed.* (YACS II) by Nuro et al. (2005) was used to outline innovative psychotherapies to reduce use of alcohol/drugs (Carroll et al., 2000; Martino, Ball, Nich, Frankforter & Carroll, 2008).

The remaining two studies identified generic fidelity measures that did not have multiple versions. The *Manual Adherence Checklist* (MAC) (Collins et al., 2009)

described a smoking cessation intervention based on either behavioural therapy or a more complex package of intervention (including behavior, cognitive, educational and interpersonal components). The MAC scale accessed for this review was in German. Prior to examining the items from this measure, each item was translated forward to English (by translator with German as a first language) and back translated by the MAC's author (Dr. Susan Collins) to verify the translation.

The *Primary Care Therapy Process Rating Scale* (PCTPRS) (Godfrey et al., 2007) was used in a study comparing cognitive-behavioural therapy with counseling sessions. Table 1 provides a summary of all selected studies, grouping together all versions of a measure.

[Insert Table 1 about here]

The five fidelity measures were utilized in nine studies on psychotherapy-based or counseling interventions for individual or family-based therapy. The measures evaluated interventions that predominantly addressed substance abuse and other mental health conditions, targeting populations from 11 to over 65 year of age. The intervention domains ranged from 2 to 4 interventions defined within the fidelity measure. All rating scales used for the generic measures were continuous 5 to 7-point Likert scales, with the exception of one dichotomous measure (Collins et al., 2009). Three of the continuous scales evaluated adherence as defined by the extensiveness of the item behaviours (i.e., frequency and intensity of performance). The remaining continuous scale examined only frequency of item behaviours (Godfrey et al., 2007). In addition to extensiveness scoring, the YACSII measure also rated items in terms of quality or competence for each item.

In terms of rating procedures, three of the measures were used to rate intervention sessions based on audio recordings (MAC; TBRS-4; PCTPRS), while the remaining two measures were used to examine audio and/or videotaped sessions. Raters scoring each session underwent training for all measures. Three of the studies used raters who were inexperienced in administering the interventions, e.g., undergraduate and/or graduate students (TBRS-4; CSPRS-6; and MAC). The other two studies used raters experienced in providing the interventions (YACSII and PCTPRS).

For the psychometric properties of each measure, the quality ratings used to rate intraclass correlations (ICC) are: <0.40 = poor; 0.40–0.59 = fair; 0.60–0.74 = good; and >0.75 = excellent (Cicchetti, 1994). Inter-rater reliability scores for the measures ranged from poor to excellent for specific domains and poor to fair for non-specific domains. Internal consistency for the subscales was reported for the TBRS-4, CSPRS-6, and PCTPRS. Cronbach's alpha values varied from fair to excellent (Cicchetti, 1994). The predominant validity tested for most fidelity measures (YACSII, TBRS-4 and CSPRS-6) was discriminant validity to differentiate between the intervention domains.

Constructs of fidelity measures

The primary conceptual foundation of generic measures focuses on evaluating the adherence of the therapist to an intervention, and differentiating therapist behaviours between interventions. To understand the other constructs of the fidelity measures, the domains and rating scales were examined. Based on the Implementation Fidelity Framework (IFF) (Carroll et al., 2007), characteristics of the measures were categorized under the constructs mediating and moderating intervention fidelity. Table 2 summarizes

the mediating and moderating constructs of each measure. The YACS II, TBRS-4 and the PCTPRS examined both mediators and moderators of fidelity. Of these measures, only the YACS II focused exclusively on therapist performance, whereas the TBRS-4 and PCTPRS evaluated both the therapists' and the clients' performance.

[Insert Table 2 about here]

All measures assessed the mediating factors of fidelity by examining therapist adherence to specific and non-specific intervention items. Three measures (YACSII, TBRS-4, and CSPRS-6) rated both frequency and intensity of the therapists' performance through one rating scale of the 'extensiveness of adherence'. The other two measures examined only the presence (MAC) or the frequency (PCTPRS) of the intervention attributes.

All measures, except for the MAC, addressed moderating attributes within the measure. Non-specific items categorized as moderators included attributes involving the client since they could impact on how the therapist delivers intervention. Two types of client attributes were observed; client behaviours and client-therapist interaction behaviours (TBRS-4, CSPRS-6 and PCTPRS). One fidelity measure (YACSII) addressed an additional moderating factor, quality of therapist delivery. To assess this attribute, all items are rated on a competence scale of the therapists' performance (i.e., therapist skillfulness to deliver the intervention).

Relevance of non-specific items to paediatric rehabilitation

The characteristics of the five generic fidelity measures were examined by two paediatric therapists to assess their applicability to a paediatric population (pre-

adolescence, age <11 years). Where possible, items have been mapped onto the FCS framework by Rosenbaum et al. (1998). Generic fidelity measures evaluate behaviours both specific and non-specific to the interventions. Since intervention-specific items can be added/deleted to any generic fidelity measure, they were not examined here. The non-specific therapist intervention items and client behaviours across the five measures in this review were examined. A total of 76 non-specific intervention items were reviewed independently by BD and ML to identify items potentially relevant to paediatric rehabilitation. This procedure yielded 91% agreement between the reviewers and a discussion about the applicability of 7 items without initial agreement resulted in 37 items (Table 3).

[Insert Table 3 about here]

The 37 non-specific items were categorized into therapist behaviours, client behaviours and client-therapist interaction behaviours. Therapist-focused items were identified in each of the five measures and were further sub-categorized into process and procedural items. Client-focused items were covered by the TBRS-4 and PCTPRS measures. Client-therapist interaction items were described by the CSPRS-6 and PCTPRS. These three categories will be described with item examples considered relevant to paediatric rehabilitation from an FCS perspective.

Therapist focused non-specific items: From the updated version of the YACS, YACS-II (Nuro et al., 2005), three items focused on therapist performance of the therapy procedure. One example of a procedure-focused item is, “To what extent did the therapist assess the patient…?” The procedure is the therapist’s performance in assessing patient

challenges or progress since last session. Ongoing assessments to monitor client progress and intervention goals are common practice in paediatric rehabilitation. The importance of assessment is evident in the FCS to identify needs and strengths for intervention (Rosenbaum et al., 1998).

One item in the MAC scale focused on the session procedure of explicitly discussing homework. A similar item was used in the CSPRS-6 (Hollon, 1984) stating, "Did the therapist review previously assigned homework with the client?" This procedure is important in the practice of FCS to provide information and to encourage participation of all family members to optimize the child's functioning (Rosenbaum et al., 1998).

Therapist non-specific items also focused on performance of the therapy process. One example item in the YACS II read, "To what extent did the therapist compliment and/or praise a past patient effort…?" Therapists showing 'confidence' and/or praising the efforts of the client takes into consideration the psychosocial needs of the child and family, as evident in the FCS framework (Rosenbaum et al., 1998). The CSPRS-6 demonstrated many process-focused behaviours with attributes such as therapist empathy, communication, rapport, encouragement, or collaboration. An example item addressing collaboration is, "Did the therapist actively attempt to engage the client in working together to explore a therapeutic issue?" Collaboration-focused items highly correspond to attributes of the FCS (Rosenbaum et al., 1998).

In the PCTPRS, non-specific items evaluated attributes related to the way therapists communicate (e.g., "Did the therapist convey warmth?"). These behaviours are

essential in FCS that emphasizes the importance of respecting and supporting families (Rosenbaum et al., 1998).

Finally, the TBRS-4 presented therapist-focused items that related to therapy process behaviours that are applicable to family therapy. For example, "(Therapist) Attempts to collaborate with parent(s) by…involving them in treatment goals." This item is consistent with the FCS that supports and encourages family involvement (Rosenbaum et al., 1998).

Client focused items: In the TBRS-4, two items examined the client's performance. These items were global in nature asking the rater "How receptive and/or engaged was the client during the session?" and "What was the level of difficulty presented in this session by the client?" Similarly, the PCTPRS (Godfrey et al., 2007) has items that examine the overall therapeutic bond in an intervention session by evaluating the behaviours of the client. One example is, the 'client works actively with the therapist's comments.' These items are directly relevant to paediatric rehabilitation, based on the importance of child engagement or participation in the child and/or the family in the FCS framework (Rosenbaum et al., 1998).

Client-therapist interaction items: The PCTPRS demonstrated non-specific items that examined therapeutic alliance between the therapist and client by evaluating the therapist-client interaction. An example of an interaction item is, "Client and therapist agree on the kind of changes to make." These are directly applicable to paediatric rehabilitation since observing the child and how he/she participates with the therapist will

indicate how well an intervention has been received and implemented (Rosenbaum et al., 1998).

Discussion

Generic fidelity measures were developed as early as the mid 1980's (Hollon, 1984) but the first appearance of the term "generic fidelity measure" in the literature was cited by Breitenstein et al. (2010). This review is the first to systematically examine generic fidelity measures and describe their characteristics. Of particular interest for paediatric rehabilitation is the potential development and application of fidelity measures that include multiple intervention domains, non-specific items, and evaluation of both therapist and client/family behaviours. Most measures utilized continuous Likert scaling consistent with the IFF concept that the rating is a 'degree' of fidelity. These measures also support observational rating by video or audio.

The value of generic measures is supported by the multiple versions of the measures evident in the review, as well as the adaptation of one measure [i.e., YACS by Carroll et al. (2000)] for a different intervention in the same area [motivational interviewing (Martino et al., 2008)]. Furthermore, multiple versions of a generic fidelity measure can exist for an intervention relevant to different populations, as seen in the CSPRS-6. Development of a generic fidelity measure applicable to paediatric rehabilitation has the potential to save time and research funds, since the development of fidelity measures remains a time-consuming and expensive process (McIntosh et al., 2005). Generic fidelity measures can be adapted to other interventions within an area of

study, so they could decrease the burden of developing a novel fidelity measure in subsequent intervention research.

All of the fidelity measures reviewed demonstrated acceptable levels of reliability and the validity analyses revealed primarily positive results. Reliability testing showed fair to excellent inter-rater reliability and internal consistency across scale-types (i.e., adherence and competence) and sub-domains (i.e., intervention specific and non-specific). A wide range of methods have been used to examine the construct validity for each measure. When considering items that are potentially adaptable to paediatric rehabilitation, new psychometric testing will be needed because one cannot assume that reliability and validity will remain constant with changes in item and population characteristics. However, the success of these generic fidelity measures for complex therapies shows promise that an analogue can be developed within the field of paediatric rehabilitation.

Almost 50% of the non-specific intervention items in the five generic measures were judged to be applicable to paediatric rehabilitation. This result was not surprising because the foundation of FCS, is based on the concept of person-centered therapy (Carl Rogers), is also integrated into many psychotherapy interventions. Most measures from the review examined behaviours that both mediated and moderated intervention fidelity. Mediating items not examined by current fidelity measures in paediatric rehabilitation include intervention non-specific items related to therapist performance. These diverse behaviours of the therapy process or procedures could form the basis of a generic measure of fidelity for pediatric rehabilitation. Some non-specific intervention items that

could be transferrable to paediatrics are "Did the therapist actively attempt to engage the client in working together …?" (Hollon, 1984); and "To what extent did the therapist compliment, or cheer, and/or praise the patient…?" (Nuro et al., 2005). Other items may provide a good starting point to investigate their relevance to paediatric rehabilitation, such as, "How involved was the therapist?" (Hogue & Samuolis, 2005).

Moderating factors are also novel characteristics for fidelity measures in paediatric rehabilitation. Several measures (YACS II, TBRS-4, and PCTPRS) included constructs that map directly onto the 'potential moderators' characteristics of the IFF. Factors that moderate fidelity in the YACSII included scale ratings of the quality of therapy delivery that evaluate therapist competence of the intervention provided, a construct shown as independent of therapist adherence. This characteristic is important to examine because therapists implementing study interventions are trained in the approaches and may vary in their proficiency. Paediatric rehabilitation interventions are multi-factored so therapist competency is important to examine since the quality of the intervention can moderate fidelity and influence outcomes.

Another attribute relevant to the construct of potential moderators in the IFF is 'client responsiveness', a concept that is also applicable to paediatric rehabilitation. Such items examined client or client-therapist interaction behaviours. Examples include, "how receptive and/or engaged was the client during the session?" (Hogue & Samuolis, 2005); and "Client works actively with therapist's comments" (Godfrey et al., 2007). In terms of client behaviours, the success of any paediatric rehabilitation intervention will most often depend on the willing involvement of the child. The impact of the responsiveness of a

child on the fidelity of the therapist's adherence needs further exploration. In addition, the importance of observing the client-therapist interaction, between the child and therapist, has been discussed in the literature in paediatric rehabilitation (Coster, Tickle-Degnen & Armenta, 1995) highlighting the need to understand the impact of this interaction on intervention fidelity. The moderating characteristics of intervention fidelity may depend on observations of the child and therapist, as well as their interactions.

Even though efforts were made to conduct a rigorous narrative review, some limitations were evident. Selected literature was limited to peer-reviewed journal articles and excluded "grey-area" sources, such as thesis dissertations. Although the exclusion of these sources demonstrates a potential selection bias to published material, the authors believe that the search was comprehensive for the current state of the field of generic fidelity measurement. Furthermore, in this work selection bias would be favourable for future work in paediatric rehabilitation, since the items from this review have reasonable psychometric properties.

There may have been fidelity measures in published research which were embedded within intervention studies and missed in the search. However, it is unlikely that these measures were generic, because most fidelity measures used in psychotherapy trials appear to be intervention-specific and are rarely used in subsequent studies (Markowitz et al., 2000). The use of a very comprehensive search strategy with the help of a health sciences librarian increases the confidence that the search strategy captured most, if not all, of the fidelity measures in the literature.

A final limitation is that study selection was determined by only one researcher (BD), raising a concern of discarding measures that may have met inclusion criteria or including measures that did not meet the inclusion criteria. The transparency of this review in describing the characteristics of each generic fidelity measure addresses the latter point, but not the former. However, most of the fidelity measures reviewed were adapted for other studies indicating that rigorous measures were included.

The next steps in the development of a generic fidelity measure for paediatric rehabilitation will be to test the list of non-specific intervention therapist attributes and client responsiveness items generated from this review. Using this list of 37 attributes, a content validation process will be conducted, followed by item development and testing for the generic fidelity measure.

Conclusion

Narrative reviews aim to "reshape previously existing information that contributes new perspectives" in an area of research (Rumrill, Fitzgerald, & Merchant, 2010, p. 399). This review has highlighted the characteristics of generic fidelity measures of health interventions applicable to interventions for children up to 11 years of age in a rehabilitation setting. From the review 76 non-specific items, 37 items were judged to be applicable to paediatric rehabilitation. This review provides foundation knowledge for research to establish a generic fidelity measure for paediatric rehabilitation.

References

Baer, J.S., Ball, S. A., Campbell, B.K., Miele, G.M., Schoener, E.P., and Tracy, K. (2007). Training and fidelity monitoring of behavioral interventions in multi-site addictions research. *Drug and Alcohol Dependence, 87,* 107–118

Bayona, C.L., McDougall, J., Tucker, M.A., Nichols, M., & Mandich, A. (2006). School-based occupational therapy for children with fine motor difficulties: Evaluating functional outcomes and fidelity of services. *Physical & Occupational Therapy in Pediatrics, 26,* 89-110.

Breitenstein, S.M., Gross, D., Garvey, C.A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation Fidelity in Community-Based Interventions. *Research in Nursing & Health, 33,* 164-173.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). *Implementation Science* 2:40 doi:10.1186/1748-5908-2-40

Carroll, K.M., Nich, C., Sifry, R.L., Nuro, K.F., Frankforter, T.L., Ball, S.A. (2000). A general system for evaluating therapist adherence and competence in psycho-therapy research in the addictions. *Drug and Alcohol Dependence, 57,* 225-238.

Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284-290.

Collins, S.E., Eck, S., Kick, E., Schroter, M., Torchalla, I., and Batra, A. (2009). Implementation of a smoking cessation treatment integrity protocol: Treatment

discriminability, potency and manual adherence. *Addictive Behaviors, 34,* 477-480.

Cook, D.J., Mulrow, C.D., and Haynes, B. (1997). Systematic Reviews: Synthesis of best evidence for clinical decisions. *Annals of Internal Medicine, 126,* 5, 376-380.

Coster, W., Tickle-Degnen, L., and Armenta, L. (1995). Therapist-child interaction during sensory integration treatment: Development and testing of a research tool. *The Occupational Therapy Journal of Research, 15,* 1, 17-35.

Diamond, G.M., Diamond, G.S., & Hogue, A. (2007). Attachment-based family therapy: Adherence and differentiation. *Journal of Marital and Family Therapy, 33,* 2, 177-191.

Godfrey, E., Chalder, T., Ridsdale, L., Seed, P., & Ogden, J. (2007). Investigating the 'active ingredients' of cognitive behaviour therapy and counseling for patients with chronic fatigue in primary care: Developing a new process measure to assess treatment fidelity and predict outcome. *British Journal of Clinical Psychology, 46,* 253-272.

Hennessey, M.L. & Rumrill, P.D. (2003). Treatment fidelity in rehabilitation research. *Journal of Vocational Rehabilitation, 19,* 123-126.

Hill, C.E. O'Grady, K., & Elkin, I. (1992). Applying the collaborative study psychotherapy rating scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. *Journal of Consulting and Clinical Psychology, 60,* 1, 73-79.

Hogue, A., and Samuolis, J. (2000). *Therapist Behavior Rating Scale – 4th Version (TBRS-4).* Fordham University, New York.

Hogue, A., Liddle, H.A., Singer, A., & Leckrone, J. (2005). Intervention fidelity in family-based prevention counseling for adolescent problem behaviors. *Journal of Community Psychology, 33,* 2, 191-211.

Hollon, S.D. (1984). Final Report: System for rating psychotherapy audiotapes. Rockville, MD: Department of Health and Human Services.

Kendall, P.C., and Beidas, R.S. (2007). Smoothing the trail for dissemination of evidence-based practices for youth: Flexibility within fidelity. *Professional Psychology: Research and Practice, 38,* 1, 13-20.

Lichstein, K.L., Riedel, B., & Grieve, R.(1994). Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research & Therapy, 16,* 1-29.

Markowitz, J.C., Spielman, L.A., Scarvalone, P.A., & Perry, S.W. (2000). Psychotherapy adherence of therapists treatment HIV-Positive patients with depressive symptoms. *Journal of Psychotherapy Practice Research, 9,* 2, 75-80.

Martino, S., Ball, S.A., Nich, C., Frankforter, T.L., & Carroll, K.M. (2008). Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence, 96,* 37-48.

McIntosh, V.V.W., Jordan, J., Mckenzie, J.M., Lutty, S.E., Carter, F.A., Carter, J.D., Frampton, C.M.A., & Joyce, P.R.(2005). Measuring therapist adherence in psychotherapy for anorexia nervosa: Scale adaptation, psychometric properties, and distinguishing psychotherapies. *Psychotherapy Research, 15,* 3, 339-344.

Moncher, F.J., & Prinz, R.J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11,* 247-266.

Morgenstern, J., & McKay, J.R. (2007). Rethinking the paradigms that inform behavioral treatment research for substance use disorders. *Addiction, 102,* 1377-1389.

Nuro, K.F., Maccarelli, L., Baker, S.M., Martino, S., Rounsaville, B.J., & Carroll, K.M. (2005). *Yale Adherence and Competency Scale - 2nd edition (YACS II).* Yale University, CT.

Parham, L.D., Cohn, E.S., Spitzer, S., Koomar, J.A., Miller, L.J., Burke, J.P. et al. (2007). Fidelity in sensory integration intervention research. *The American Journal of Occupational Therapy, 61,* 216-227.

Roberts, H. (2007). *Systematic Reviews In the Social Sciences: A practical guide.* Oxford: Wiley-Blackwell.

Rosenbaum, P., King, S., Law, M., King, G., and Evans, J. (1998). Family-Centred Service: A conceptual framework and research review. *Physical & Occupational Therapy in Pediatrics, 18,* 1, 1-20.

Rumrill, P.D., Fitzgerald, S.M., and Merchant, W.R. (2010). Using scoping literature reviews as a means of understanding and interpreting existing literature. *Work, 35,* 399-404.

Smith, S.W., Daunic, A.P., Taylor, G.G. (2007). Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. *Education and Treatment of Children, 30,* 4, 121-134.

Waltz, J., Addis, M.E., Koerner, K. & Jacobson, N.S. (1993). Testing the integrity of psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and clinical psychology, 61,* 620-630.

Waskow, I.E. (1984). Specification of the technique variable in the NIMH Treatment of Depression Collaborative Research Program. In: Williams, J.B.W., Spitzer, R.L. (Eds.), Psychotherapy Research: Where are we, and where should we go? Guilford, New York.

Whyte, J. (2003).Clinical trials in rehabilitation: What are the obstacles? *American Journal of Physical Medicine and Rehabilitation, 82,* 10(Suppl), S16-S21.

Table 1: Summaries of the characteristics of selected fidelity measures

Author (Year)	Fidelity Measure	Population (age category)	Intervention Domains	Rating Scale(s)	Psychometric Testing
Therapeutic Behavior Rating Scale (TBRS)					
Diamond et al. (2007)	Therapeutic Behavior Rating Scale 3rd version (TBRS-3)	Depression or substance users (adolescents and adult family)	**Family therapy:** ABFT clients with depression; MDFT for substance users; & CBT.	7-point Likert scale, scoring extensiveness item (frequency and intensity)	**IRR**: intervention domains ICC (.69 to .96); non-specific domain (.54 to .67). **IC:** α >.50 between item Correlations. **Discriminant Validity:** interventions were independent and differentiable.
Hogue et al. (2005)	Therapeutic Behavior Rating Scale 2nd version (TBRS-2)	High risk for substance abuse and related issues (adolescent, and adult family)	**Family-based preventative intervention:** MDFP; MDFT for at risk adolescents; & CBT.	7-point Likert scale, scoring extensiveness item (frequency and intensity)	**IRR:** Interventions ICC (.71to .84). **IC**: α ranged from .43 to .74 **Validity:** Showed predictable convergence and discrimination validity between sub-scales.
Yale Adherence and Competence Scale (YACS)					
Martino et al. (2008)	Independent Tape Rating Scale (ITRS) *(adaptation of YACS)*	Outpatients with mixed substance use issues (adults)	**Motivational interview Drug counseling therapy:** MET; & Usual Counseling	7-point Likert scoring item extensiveness (frequency and intensity) and therapist competence	**IRR:** ICC for adherence (.55 to .98) & competence (.67 to .98) scales. **Discriminant validity**: Evident between interventions based on factor structure.

Carroll et al. (2000)	Yale Adherence and Competence Scale (YACS)	Drug use disorders (adults)	**Behavioral Therapy:** CM; TSF; CBT	5-point Likert scoring item extensiveness (frequency and intensity) and therapist competence	**IRR:** ICC for adherence (.80 to .95) and for competence (.71 to .98). **Validity:** Excellent factor structure, construct validity for sub-domains and between dimension of adherence and competence; Concurrent and Discriminant validity demonstrated.
Collaboration Study Psychotherapy Rating Scale (CSPRS)					
McIntosh et al. (2005)	Collaborative Study Psychotherapy Rating Scale for Anorexia Nervosa (CSPRS-AN)	Anorexia Nervosa Outpatient (adolescents and adults)	**Therapy:** CBT; IPT; & NSCM (Usual practice).	7-point Likert scale, scoring extensiveness item (frequency and intensity)	**IRR:** for each intervention (.86 to .94), but (.39) for non-specific scale. **IC:** for all subscales $\alpha \geq .82$
Markowitz et al. (2000)	Collaborative Study Psychotherapy Rating Scale for HIV (CSPRS-HIV)	HIV-positive depressed outpatients (not provided)	**Therapy:** CBT; IPT; SP; & SWI.	7-point Likert scale, scoring extensiveness item (frequency and intensity)	**IRR**: for each specific and non-specific intervention (.89 to .99) domains. **Discriminant validity:** across 4 interventions using adherence scale in ANOVA.
Hill et al. (1992)	Collaborative Study Psychotherapy Rating Scale (CSPRS)	Adults with Depression (adult)	**Therapy:** CBT; IPT; Imipramine; & Pill placebo with Clinical Management.	7-point Likert scale, scoring extensiveness item (frequency and intensity)	**IRR**: for each specific intervention (.58 to .88) and for non-specific domain (0.47). **IC**: α = .5 to .86 **Discriminant validity:** demonstrated for interventions involved.
Manual Adherence Checklist (MAC)					
Collins et al. (2009)	Manual Adherence Checklist	Cigarette Smokers (adult)	**Smoking Cessation Counseling:**	-Entire session coded with dichotomous	**IRR**: ranged from .24 to perfect agreement (κ =1.00) was achieved for 53% of the treatment components.

	(MAC)		BT "Standard"; BT with cognitive, psycho-educational and interpersonal components.	scoring. [coding system by Waltz et al. (1993)]	**Validity:** Interventions not highly discriminable.
Primary Care Therapy Process Rating Scale (PCTPRS)					
Godfrey et al. (2007)	Primary Care Therapy Process Rating Scale (PCTPRS)	Chronic Fatigue in Primary Care (adolescents, adults, and older adults)	**Therapy:** CBT; & Counseling	7-point Likert scale, scoring adherence of performance (i.e., frequency)	**IRR:** for subscales ranged from poor to excellent. **IC**: for all subscales α = excellent. **Validity:** Face validity of therapy using each scale.

Note. ABFT = Attachment-Based family therapy; MDFT = Multidimensional family therapy; CBT = Cognitive-Based Therapy; MDFP = Multidimensional family prevention; MET= Motivational enhancement therapy; CM = Clinical Management; TSF = Twelve Step Facilitation; IPT = Interpersonal Therapy; NSCM = Nonspecific Supportive Clinical Management (Usual practice); SP = Supportive Psychotherapy; SWI = Supportive psychotherapy with imipramine; BT = Behavioral Treatment; IRR = Inter-rater Reliability; IC = Internal Consistency

Table 2: Conceptual foundations of fidelity measures based on item domains and scaling

Measure (authors)	Mediators of intervention Evaluated	Potential Moderators Evaluated
TBRS – 4th Edition (Hogue & Samuolis, 2000)	Adherence of **therapist** performance of intervention **specific and non-specific items** [Scale: Extensiveness (frequency and intensity)]	Non-Specific Items examining **client behaviours** and **client-therapist interactions**.
YACS II (Nuro et al., 2005)	Adherence of **therapist** performance of intervention **specific and non-specific items** [Scale: Extensiveness (frequency and intensity)]	Rating scale scoring the **quality of therapist performance** of specific and non-specific items.
CSPRS-6 (Hollon, 1984)	Adherence of **therapist** performance of intervention **specific and non-specific items** [Scale: Extensiveness (frequency and intensity)]	Non-Specific items examining **client-therapist interactions**.
MAC (Collins et al., 2009)	Adherence of intervention content from **therapist** performance for **specific and non-specific items**. [Scale: presence or absence of item behaviour]	None
PCTPRS (Godfrey et al., 2007)	Adherence of **therapist** performance of intervention **specific and non-specific items** [Scale: Frequency of item delivery]	Non-Specific Items examining **client behaviours** and **client-therapist interactions**.

Table 3: Non-specific items from fidelity measures relevant to paediatric rehabilitation

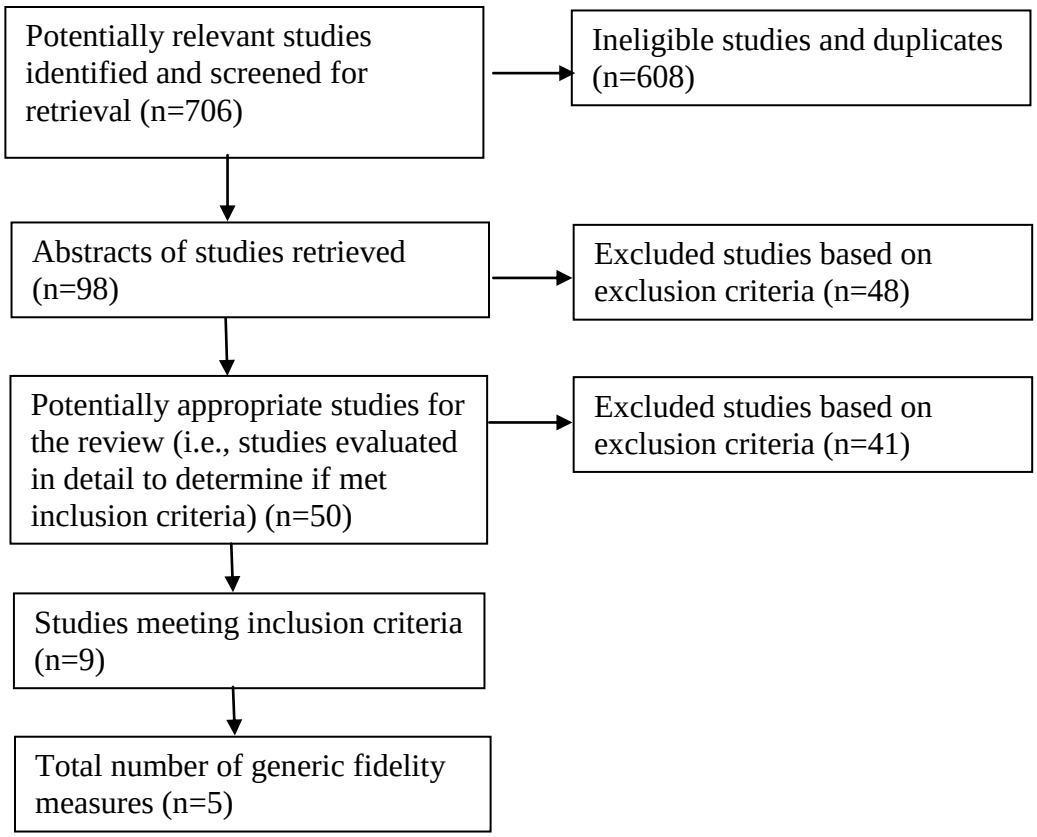
Item	Item source
Therapist behaviours	
1. TO WHAT EXTENT DID THE THERAPIST ASSESS THE PATIENT'S use of 'primary drug' SINCE THE LAST SESSION, including the pattern of substance use (if any)?	YACSII
2. TO WHAT EXTENT DID THE THERAPIST EXPLICITLY FOCUS ON THE PATIENT'S PSYCHOPATHOLOGY (i.e. symptoms of depressive, anxiety, psychotic disorders)?	YACSII
3. TO WHAT EXTENT DID THE THERAPIST ASSESS THE PATIENT'S GENERAL LEVEL OF FUNCTIONING in major life spheres (e.g., work, intimate relationships, family life, social life, everyday stress, etc.)?	YACSII
4. TO WHAT EXTENT DID THE THERAPIST COMPLIMENT AND/OR PRAISE A PATIENT EFFORT that did not include the role of *medication*?	YACSII
5. TO WHAT EXTENT DID THE THERAPIST INQUIRE ABOUT OR DISCUSS THE AVAILABILITY AND NATURE OF SUPPORT FROM FAMILY MEMBERS and/or significant others for the patient's involvement in treatment or efforts to become abstinent?	YACSII
6. TO WHAT EXTENT DID THE THERAPIST COMMUNICATE CONFIDENCE THAT PATIENT EFFORTS that do not include medication WILL YIELD SUCCESS IN THE FUTURE?	YACSII
7. TO WHAT EXTENT DID THE THERAPIST COMPLIMENT, OR CHEER, AND/OR PRAISE THE PATIENT'S REPORTED *study medication* COMPLIANCE?	YACSII
8. TO WHAT EXTENT DID THE THERAPIST DISCUSS, REVIEW, OR REFORMULATE THE PATIENT'S GOALS FOR TREATMENT?	YACSII
9. DID THE THERAPIST WORK COLLABORATIVELY WITH THE CLIENT TO FORMULATE AND FOLLOW A SPECIFIC AGENDA FOR THE SESSION?	CSPRS-6
10. DID THE THERAPIST REVIEW PREVIOUSLY ASSIGNED HOMEWORK WITH THE CLIENT?	CSPRS-6
11. WAS THE THERAPIST SUPPORTIVE OF THE CLIENT BY ACKNOWLEDGING THE CLIENT'S GAINS DURING THERAPY or by reassuring the client that gains will be forthcoming?	CSPRS-6
12. DID THE THERAPIST CONVEY THAT HE/SHE UNDERSTOOD THE CLIENT'S PROBLEMS and is ABLE TO HELP THE CLIENT?	CSPRS-6

13. DID THE THERAPIST CONVEY WARMTH?	CSPRS-6
14. WAS THE THERAPIST EMPATHETIC TOWARD THE CLIENT (i.e., did she/he convey an intimate understanding of and sensitivity to the client's experiences and feelings)?	CSPRS-6
15. DID THE THERAPIST ACTIVELY ATTEMPT TO ENGAGE THE CLIENT IN WORKING TOGETHER to explore therapeutic issues?	CSPRS-6
16. DID THE THERAPIST ATTEMPT TO TEACH THE CLIENT SKILLS IN THE SESSION?	CSPRS-6
17. DID THE THERAPIST OR CLIENT DEVELOP ONE OR MORE SPECIFIC ASSIGNMENTS FOR THE CLIENT TO EXECUTE BETWEEN SESSIONS?	CSPRS-6
18. DID THE THERAPIST ENCOURAGE THE CLIENT TO ENGAGE IN ACTIVITIES WHICH WOULD BE PLEASURABLE TO THE CLIENT or from which the client would obtain a sense of accomplishment?	CSPRS-6
19. DID THE THERAPIST ENCOURAGE THE CLIENT TO ENGAGE IN ACTIVITIES WHICH WOULD BE PLEASUREABLE TO THE CLIENT OR FROM WHICH TH CLIENT WOULD OBTAIN A SENSE OF ACCOMPLISHMENT?	CSPRS-6
20. (The therapist) ATTEMPTS TO ESTABLISH AND MAINTAIN the adolescent's INVESTMENT IN THERAPY AND/OR FORMULATE TREATMENT GOALS	TBRS-4
21. (The therapist) ATTEMPTS TO COLLABORATE WITH PARENT(S) BY INSTILLING HOPE AND/OR INVOLVING THEM IN TREATMENT GOALS	TBRS-4
22. (The therapist) BUILDS OR EXPLORES CONNECTIONS BETWEEN PARENT(S) and the adolescent's ecosystem	TBRS-4
23. (The therapist) PRESENTS KNOWLEDGE ABOUT NORMATIVE adolescent DEVELOPMENT	TBRS-4
24. (The therapist) ARRANGES, COACHES, AND PROCESSES MULTIPARTICIPANT INTERACTIONS IN SESSION	TBRS-4
25. (The therapist) FOCUS ON PARENT'S NON-PARENTING LIFE AS AN ADULT PERSON	TBRS-4
26. ARRANGES, COACHES, AND PROCESSES MULTI-PARTICIPANT INTERACTIONS IN SESSION.	TBRS-4
27. FOCUS ON PARENT'S NON-PARENTING LIFE AS AN ADULT PARENT	TBRS-4
28. TARGETS PARTICIPANTS OTHER THAN THE ADOLESCENT FOR CHANGE.	TBRS-4
29. HOW INVOLVED WAS THE THERAPIST?	PCTPRS
Client Behaviours	
30. HOW RECEPTIVE AND/OR ENGAGED WAS THE CLIENT DURING THE SESSION?	TBRS-4

31. WHAT IS THE LEVEL OF DIFFICULTY PRESENTED IN THIS SESSION BY THE CLIENT?	TBRS-4
32. CLIENT WORKS ACTIVELY WITH THE THERAPIST'S COMMENTS	PCTPRS
33. CLIENT SHOWS CONFIDENCE IN THERAPY AND THERAPIST	PCTPRS
Client-therapist interaction behaviours	
34. HOW MUCH RAPPORT WAS THERE BETWEEN THERAPIST AND CLIENT (i.e. how well did the therapist and client get along)?	PCTPRS
35. CLIENT AND THERAPIST AGREE ON THE KIND OF CHANGES TO MAKE	PCTPRS
36. CLIENT AND THERAPIST SHARE SAME SENSE ABOUT HOW TO PROCEED	PCTPRS
37. CLIENT AND THERAPIST AGREE on salient themes	PCTPRS

Note. CSPRS-6 = Collaborative Study Psychotherapy Rating Scale - 6th Version (Hollon, 1984); YACS II = Yale Adherence and Competence Scale - 2nd Ed. (Nuro et al., 2005); TBRS-4 = Therapist Behavior Rating Scale – 4th Version (Hogue and Samuolis, 2000); PCTPRS = Primary Care Therapy Process Rating Scale (Godfrey et al., 2007).

Figure 1: Flow chart of study selection in primary search



Chapter Three

General Therapy Behaviours in Paediatric Rehabilitation: Essential Attributes for Intervention with Children with Physical Disabilities

Authors: Briano Di Rezze, Mary Law, Kevin Eva, Nancy Pollock, Jan Willem Gorter

This chapter contains a manuscript entitled: "General therapy behaviours in paediatric rehabilitation: essential attributes for intervention with children with physical disabilities". The proposed journal for this manuscript is Disability and Rehabilitation. This manuscript is currently in an expanded format, and is yet to be submitted to Disability to Rehabilitation.

Abstract

Paediatric rehabilitation is a dynamic process that involves the therapist delivering intervention-specific and non-specific behaviours to help the client achieve his/her goals. Non-specific intervention characteristics of the therapy process, including interpersonal relationships, are a key part of family-centred service (FCS). Despite its importance, there is little research that identifies observable therapeutic behaviours that can be used to identify the fidelity of intervention to FCS principles and to examine their impact on outcomes. **Objective:** To generate observable general therapy attributes of paediatric rehabilitation essential to family-centred rehabilitation interventions for children with physical disabilities. **Methods:** Attributes of general therapy behaviours were derived based on a Delphi Process with multidisciplinary researchers and rehabilitation therapists. A separate method identified attributes through the content analysis of semi-structured interviews with practicing occupational therapists and physiotherapists. These data

underwent a triangulation procedure to corroborate general therapy behaviours relevant to family-centred service. **Results/Discussion:** Eight researchers participated in the Delphi Process. Seventeen therapists participated in semi-structured interviews. The Delphi Process generated 35 behavioural attributes. These attributes were divided into the following categories: therapist behaviours (21); client behaviours (9); and client-therapist behaviours (5). Of the 19 attributes generated from the content analysis of therapist interviews, 17 mapped onto the attributes identified in the Delphi Process. The codes that did not match Delphi attributes related to therapist interactions with other health care professionals, rather than focusing on concepts directly related to FCS. These general therapy attributes addressed a range of behaviours including characteristics of the intervention procedure (e.g., assessment) and the therapeutic process involving both the child and parent with the therapist. This work provides an improved understanding of how practitioners conceive essential and observable behaviours of FCS intervention that will enable future researchers to identify their presence within an intervention session.

3-5 points of what the paper adds: 1) This paper broadens the focus of fidelity measurement of paediatric rehabilitation to include behaviours relevant to family-centred service. 2) Attributes of the therapist's general practice behaviour in family-centred service were identified. 3) Attributes of paediatric rehabilitation involving the child's response to intervention, parent participation and child and therapist interaction were generated.

Introduction

Emerging methods to improve the rigour of paediatric rehabilitation research focus on defining the active ingredients of interventions and establishing methods to evaluate therapist adherence to intervention delivery (Whyte & Hart, 2003; Nelson & Mathiowetz, 2004). Together, these methods measure how well a therapist delivers an intervention as planned, a concept known as intervention fidelity (Carroll et al., 2007). Measurement of intervention fidelity enables investigators to explore the relationships between interventions and outcomes more rigorously.

In addition to examining intervention-specific strategies, there has been discussion about other active ingredients of therapeutic intervention that potentially impact outcomes (Abraham & Michie, 2008). Non-specific intervention behaviours commonly employed across therapies may also contribute to changes in human behaviours. These general types of therapy behaviours can take the form of common therapy procedures (e.g., conducting an assessment); behaviour change strategies (e.g., modeling behaviour or providing feedback on performance) (Abraham & Michie, 2008); or aspects of the therapy process (or help-giving behaviours) within family-centred practice (e.g., active listening or empathy) (Dunst, 2009). Even though there is no literature in paediatric rehabilitation identifying the impact of these non-specific (or general) therapy behaviours on outcomes, some behaviour change approaches within adult intervention have shown to be associated with effectiveness (Albarracin et al., 2005).

In addition, areas in psychology have determined that such strategies (e.g., therapeutic alliance) can influence outcomes (Elvins & Green, 2008). Thus, in the examination of the fidelity of an intervention, the identification of all key influences within an intervention session on a child's outcome would include these general characteristics of therapy. All non-specific intervention characteristics (including therapy process, behaviour change, and therapy procedures) will be referred to as general therapy behaviours in this paper.

In the paediatric rehabilitation literature, family centered service (FCS) delivery is considered best practice (Rosenbaum et al., 1998; Dunn, 2011). Attributes of general therapy behaviours can influence family-centred service delivery (Dickens, Matthews, and Thompson, 2011; Law et al., 2003). Specifically, interpersonal attributes of the therapeutic process are important to families (Law et al., 2003) and are associated with greater client satisfaction (King, King & Rosenbaum, 1996). Work by Dunst and colleagues (2006) identified two dimensions of help-giving that are provided to families in FCS: relational help-giving, such as empathy and listening skills; and participatory help-giving, involving more action oriented behaviours such as individualized or flexible practice. Currently, no research has examined observable general therapy behaviours reflecting these dimensions in the context of intervention sessions that follow a FCS philosophy. Because this area is under-researched in children and adolescents (Faw, Hogue, & Liddle, 2005), there are no validated methods to observe and assess essential general therapy behaviours as part of measuring intervention fidelity. The first step in being able to evaluate the impact of such characteristics within paediatric rehabilitation would be to identify general therapy behaviours. The purpose of this study was to

generate observable attributes of general therapy behaviours that are considered to be essential within rehabilitation intervention sessions for children with physical disabilities within a family-centred services framework.

Background

Trends in examining complex therapeutic intervention

To improve the rigour of rehabilitation research, it is necessary to define and measure the fidelity of delivering the active ingredients of interventions (Nelson & Mathiowetz, 2004). Detailed descriptions of these characteristics can increase the understanding of the mechanisms of intervention as well as increase confidence that the intervention was implemented as planned.

Measures of intervention fidelity typically evaluate the degree and quality of therapist performance of the specific intervention under investigation. Core competencies of health care delivery across disciplines include clinical skills for specific interventions as well as interpersonal skills (Verma et al., 2006). In the adult psychotherapy literature, researchers have recognized the need to also examine general therapy behaviours through generic fidelity measures. Generic fidelity measures often examine therapy attributes common across interventions alongside specific fidelity measures to differentiate unique interventions (Breitenstein et al., 2010; Di Rezze, Law, Gorter, Eva, & Pollock, In Press). A review of common generic attributes showed that they focus on general therapy behaviours, including behaviours of the therapist (e.g., praising effort), the client (e.g., responsiveness) and the interaction behaviours between the therapist and client (e.g.,

collaboration) within an intervention session (Di Rezze et al., In Press). Although the generic measures in the review were predominantly adult focused, general therapy behaviours are also important characteristics of paediatric rehabilitation.

Paediatric Rehabilitation Intervention

Paediatric rehabilitation interventions for children with disabilities are delivered by multidisciplinary teams involving both physicians (e.g., developmental paediatricians) and health professionals (e.g., occupational therapists, physiotherapists) (Hong & Palmer, 2003). Best practice in paediatric rehabilitation is based on the principles of family-centred service (Rosenbaum et al., 1998; Dunn, 2011). Although several terms for family-centred service exist in the literature (e.g., family-centred practice or care), all indicate that specific elements of general therapy behaviours between the child, parent and therapist facilitate optimal intervention outcomes (Green, 2009). In a systematic review examining experiences of family-centred care for children in hospital, important characteristics to operationalize family-centred care included communication and the relationship between the health provider and parent (Foster, Whitehead, & Maybee, 2010). Another review concluded that the interpersonal aspects of family-centred health care practice, such as information exchange, respectful and supportive care and partnership, were associated with outcomes of greater client satisfaction (King et al., 1996).

A wide range of principles within general therapy behaviours in family-centred practice have been specified. Only self-report measures have been developed in this area for therapists (Dunst et al., 2006; King et al., 1996). However, there is a need to identify a

comprehensive list of observable behaviours within an intervention session that could potentially have an impact on outcomes.

Methods

This study used two data sources to identify attributes of general therapy behaviours essential within an intervention session for children with physical disabilities. Through methods of triangulation, the main source of data (core component) was supplemented by additional data (supplemental component) to ensure a complete description of the investigated phenomenon (Junious et al., 2010). Core component data were generated through a Delphi Process with rehabilitation researchers and supplemented by qualitative interviews with paediatric therapists.

The Delphi Process asked multidisciplinary researchers in the field of childhood disability to identify and rate the applicability of general therapy behaviours. This procedure determines the extent to which experts achieve consensus about a specific issue using sequential multistage survey rounds in a short timeframe (Jones & Hunter, 1995; McKenna, 1994). Findings from the Delphi Process were supplemented with attributes identified through a content analysis of semi-structured qualitative interviews with paediatric physical and occupational therapists. The interviews with paediatric therapists were part of a larger study that received approval from the Faculty of Health Sciences research ethics board (REB) at McMaster University. Separate approval for the Delphi Process was received from this REB.

Generation of Core Component Data

Delphi Process Procedure

A purposeful sample (Morse & Richards, 2002) of experienced clinician-researchers in paediatric rehabilitation was targeted for the Delphi process. Recommended Delphi Process procedures were followed for this study and participants were informed that a minimum of two survey rounds were planned to achieve consensus (McKenna, 1994). New surveys were generated for each round. To maximize survey response rates, strategic principles of the Tailored Design Method (TDM) for survey administration and survey design were followed (Dillman, 2007).

The Delphi surveys were web-based, because this was a familiar medium for the participants and an efficient way to receive prompt responses. Time to complete each survey was approximately 30 minutes and participants were requested to complete each survey within two weeks. Participants were sent reminders when one week and one day remained for each round (Dillman, 2007). All surveys were developed and revised between rounds to minimize redundancy by the first author (BD), and then reviewed by his supervisor (ML) to ensure clarity of items.

The Delphi Process started with a list of general therapy behaviours identified as potentially relevant to paediatric rehabilitation during a narrative review of generic fidelity measures (Di Rezze et al., In Press)**.** Respondents were asked to rate the relevance and clarity of each item as an essential behaviour in an intervention session for paediatric rehabilitation. Ratings for attributes were made on a 4-point Likert scale (1 = not relevant to 4 = very relevant and succinct) (Lynn, 1986). To establish consensus

agreement using this scale, scores were dichotomized whereby scores of 3 and 4 demonstrated endorsement of the item and scores of 1 and 2 did not endorse the item. Since the Delphi Process seeks consensus and does not resolve disagreement (McKenna, 1994), attributes not achieving consensus over two rounds were eliminated from the pool. Consensus methods for analyzing participant responses and establishing a priori standards for consensus were based on standards established by Lynn (1986). To conclude that consensus and content validity for each attribute was achieved beyond the $p<0.05$ level of significance, a minimum agreement was required to endorse each attribute as per the calculations performed by Lynn (1986).

In Round 1, all 35 potential attributes were grouped into three categories, Therapist, Client and Client-Therapist interaction. Following the attribute ratings of each category, participants were asked to rate the completeness of that category on a 4-point scale (1= list of items almost fully incomplete; 4=list of items almost fully complete). For all three rounds of the Delphi Process, participants were provided with opportunities to suggest additional attributes that were not identified within each category.

After Round 1, attributes were sorted into two groups, consensus achieved or not achieved. In Round 2, experts re-rated attributes (with original wording) from Round 1 that lacked consensus; previous scores were provided to indicate the dispersion of scores across raters (Jones & Hunter, 1995). Participants were asked to consider the relevance of core content as essential to any intervention session for paediatric rehabilitation. The survey in Round 2 also requested ratings of the relevance of new attributes suggested by

respondents in Round 1. Duplicate attribute suggestions from Round 1 were eliminated and similar concepts were combined by BD.

The survey in Round 3 asked participants to re-rate attributes introduced for the first time in Round 2 that did not achieve consensus. Attributes that received consensus in the two previous rounds were re-phrased to apply directly to paediatric rehabilitation and were re-rated in terms of their relevance. This process generated scores of agreement about relevance based on the final wording of all attributes. Round 3 concluded with respondents providing ratings on the comprehensiveness of the final list of attributes for each category (i.e., Therapist, Client and Client-Therapist). See Figure 1 for a flow chart outlining the Delphi procedure across three rounds.

[Insert Figure 1 about here]

Generation of Supplemental Component Data

In addition to the Delphi Process, an independent sample of rehabilitation therapists participated in semi-structured interviews in which they were asked about general attributes essential to best-practice paediatric intervention. The interviews were conducted over telephone by BD with a convenience sample of occupational therapists (OT) and physiotherapists (PT) as part of a study to examine their experience of conducting an intervention within a randomized trial (Law et al., 2011). The interviews were audio-recorded and professionally transcribed. These data provided a supplementary source of general attributes of rehabilitation intervention that were independent of those from the Delphi Process.

Content Analysis Procedure

Interview questions focused on identifying the essential attributes of any paediatric intervention. Interview responses underwent descriptive content analysis, which "does not require the researcher to move as far from the data" (Sandelowski, 2000, p.335). Descriptive content analysis was utilized to focus on statements based on surface meaning or everyday language. Content was considered "manifest content" when it described visible and obvious components. The content unit was the constellation of statements or sentences in the transcripts with the same central meaning (Graneheim & Lundman, 2004).

Content analysis involved a systematic process of labeling the data units (and identifying their meaning), establishing a reliable coding scheme, and applying the coding scheme to all data (Ezzy, 2002). The transcripts were labeled by BD and memos (Morse & Richards, 2002) were made about the meaning of the data units. The labeled units were reviewed to identify similar codes and finalized by defining each code (Hsieh & Shannon, 2005). The internal validity of the coding scheme was tested by randomly selecting and coding two transcripts (Elo & Kyngas, 2008) between BD and an independent rater (paediatric occupational therapist). The criterion for agreement was a minimum of 75%. Data units which fell below 75% were resolved by consensus and coding was retested with two new randomly selected transcripts. The coding scheme was then systematically applied to all interview transcripts.

Data Analysis

Descriptive analyses were used to present participant characteristics for the Delphi Process and therapist interviews. For the Delphi Process data, frequency

distributions were used to identify patterns of consensus for the relevance of attributes, as well as the completeness of the attributes listed for each category (i.e., therapist, client, client-therapist interactions). Descriptive content analysis data were presented as frequency distributions for the codes across transcripts.

To combine the attributes generated from the Delphi Process and the supplemental content analysis, a systematic triangulation procedure was determined a priori. Attributes generated from the core and supplemental data sources were tabled (Adami & Kiger, 2005). An additional table column paired similar attributes and noted the similarities or differences between the paired attributes (Casey & Murphy, 2009). These data were re-examined by another researcher (ML); any disagreements were discussed and decided by consensus. Unmatched attributes and codes were not incorporated into the final list of attributes.

Results

Delphi Process

For the Delphi Process, eight health researchers in the area of childhood physical disability met the inclusion criteria and consented to participate. These participants were recruited via an internationally recognized childhood disability research centre. Members of this research centre were knowledgeable about the practice and study of family-centred rehabilitation intervention for children with physical disabilities in North America. Their professional background included developmental paediatrics (n=2), kinesiology (n=1), occupational therapy (n=3), and physiotherapy (n=2). All experts had over 11 years of

experience working in paediatric rehabilitation and all worked with children from age 3-12 years. All respondents had experience working with children with cerebral palsy and the majority (n=5) of respondents also had experience with children with other motor and developmental disabilities. All respondents (n=8) participated in Rounds 1 and 2, while one dropped out in Round 3.

Thirty-five potential attributes were rated by all eight participants in Round 1. Based on the work by Lynn (1986), a minimum of 7/8 (88%) participants were required for consensus to endorse (scores 3 or 4) or eliminate (scores 1 or 2) each attribute. Table 1 provides a detailed summary of the attributes and their status after each round of rating. In Round 1, consensus was reached to keep 24 attributes. The remaining 11 attributes did not achieve consensus and were re-rated in Round 2. Based on respondent comments, 26 new potential attributes were added after Round 1.

[Insert Table 1 about here]

As a result, 37 attributes were rated in Round 2 (11 re-rated and 26 new). Of the 11 re-rated attributes, five achieved consensus to be kept and 6 were eliminated. Twenty two of the new attributes achieved consensus and were kept. The remaining four attributes did not achieve consensus and were re-rated in Round 3.

No new items were generated in Round 2. Following Rounds 1 and 2, there were 54 potentially relevant attributes. These attributes were reviewed by BD and ML to: (a) identify redundancies; (b) simplify concepts by dividing them into more than one attribute; and (c) re-word the final list of attributes to be directly applicable to paediatric rehabilitation. After review, 21 were not changed, 32 attributes were collapsed into 13,

and one was divided into two. As a result, 36 attributes were rated in Round 3, with all attributes achieving consensus (35 to keep and 1 to eliminate). At least 86% of respondents rated each of the three categories to be a complete list of attributes. The final number of attributes for each category were (35): therapist behaviours (21); client behaviours (9); and client-therapist behaviours (5). See Table 2 for the final list of attributes generated through the Delphi Process.

[Insert Table 2 about here]

The 7 eliminated attributes were primarily therapist behaviours (6). These attributes generally focused on the therapist's level of involvement and playfulness, in addition to their level of attention to other participants within the intervention session. One eliminated attribute focused on client behaviour, specifically the level of difficulty the client presented within the session.

Therapist Interviews

Twenty therapists were contacted via electronic mail to participate in the semi-structured interview and 17 consented (8 occupational therapists and 9 physiotherapists). Nine therapists had over 10 years of experience in paediatric rehabilitation. For the remaining therapists, four had over 5 years of experience and three therapists had between 1-5 years of experience. Demographic data were missing for one therapist.

Coding reached 75% agreement between BD and an independent rater after two sets of transcript dyads were co-rated independently. The coding scheme included 19 codes in 10 categories. The ten code categories were: Therapist-parent partnership; Therapist-parent rapport; Parent education; Motivating the child; Therapist-child rapport;

Empowering the child; Strategies to facilitate child's learning; Session procedures; Therapist general characteristics; and Therapist's interdisciplinary interactions. See Table 3 for examples of codes within several sample categories.

[Insert Table 3 about here]

From these 19 codes, 17 were similar and two were different from the attributes in the Delphi Process. See Table 2 for attributes and codes mapped on where applicable. From the 17 codes, 2 mapped onto the same attribute, resulting in 15 attributes possessing similar content to the codes. The definitions from the coding scheme provide more detail on the observable behaviour of these 15 attributes. For example, the attribute labeled *Therapist compliments and/or praises the child's effort or performance within the intervention session* was defined by the coding scheme as a context whereby, "the therapist creates a positive learning environment for the child based on their observed performance in the intervention session by celebrating or rewarding achievements in session". Definitions of the remaining attributes can be provided by the primary author upon request. The two novel attributes that emerged from the content analysis related to collaboration with other health professionals and was in the therapist behaviour category, e.g., therapist collaborates on agenda within interdisciplinary session.

Discussion

Over the last 20 years, the principles of Family-Centred Service have been clearly described (Rosenbaum et al., 1998). As well, self-report of practitioner service behaviours can be completed through existing measures such as the Measure of Process

of Care (MPOC) (King, Rosenbaum, & King, 1995). The MPOC identifies service level behaviours that are not measuring particular behaviours from an intervention session. Thus, what is lacking is the identification of specific and observable behaviours within the therapeutic process that should be present in any paediatric rehabilitation intervention that adheres to FCS principles. From the 35 attributes identified in the Delphi study, 15 overlapped with content from the therapist interviews. The majority of these overlapped attributes (8) were concepts of participatory help-giving (Dunst, 2009), e.g., *THERAPIST alters his/her behaviour to adapt to the challenges, needs, behaviours, interests of the CHILD in the intervention session*. The remaining 7 attributes aligned with concepts of behaviour change (e.g., praising the client) (Abraham & Michie, 2008) and relational help-giving (e.g., being respectful) (Dunst, 2009). Hence, these attributes align with some concepts identified generally in the MPOC and from the FCS literature as important to families, such as respect, collaboration, flexibility in therapy services, interpersonal characteristics, and support (Dickens, et al., 2011; Law et al., 2003; Rosenbaum et al., 1998).

The behaviour change strategies identified as essential attributes in paediatric rehabilitation increase the scope of common behaviours to be observed and evaluate within interventions. These general therapy behaviour attributes represent a broader view of what is important in FCS through the identification of behaviours of all persons involved in paediatric interventions, including the child, parent, therapist and their interactions. Attributes that were eliminated throughout the Delphi process defined

practice behaviours narrowly (e.g. therapist being playful) or were unrelated to paediatric rehabilitation (e.g., psychopathology of the client).

The results of this study demonstrated a high degree of consensus reached by a multidisciplinary group of paediatric health professionals on a wide range of attributes across participants within an intervention session (i.e., clients and therapists). The attributes reaching agreement in the Delphi Process were a combination of those derived from the psychotherapy literature and the paediatric rehabilitation professionals, amalgamated into 35 attributes. Items derived from the psychotherapy literature were expected to be relevant to paediatric rehabilitation, since FCS principles are based on the concept of client-centred practice, a foundation of psychotherapy practice derived by Carl Rogers. Most of the attributes eliminated from the Delphi process were therapist behaviours that appear to be outside of the scope of paediatric rehabilitation (e.g., focus on psychopathology of the client, manipulating the interactions of the participants in the session, or focusing on non-parenting life) or too specific in defining therapist behaviours (e.g., assessment since last session, how involved was therapist, and was therapist playful).

This study derived additional content relevant to the FCS model by focusing on the family's involvement in rehabilitation sessions. Elements of the parent-therapist relationship and other characteristics of client-centred relationship that are child specific can influence intervention sessions. These results are consistent with emerging paediatric literature regarding the therapeutic alliance as described by Green (2009).

This study employed a rigorous triangulation process that included experienced participants in paediatric rehabilitation research and clinical practice. The attributes generated from the Delphi Process represented a near complete list that were validated from practicing therapists, since 17 of 19 therapist-interview codes mapped onto Delphi derived attributes. The two codes that did not map onto the Delphi attributes were related to working with colleagues in a multidisciplinary practice setting. This practice context is not always how paediatric rehabilitation is conducted, but can be considered as an essential attribute in multidisciplinary teams. Including these multidisciplinary attributes, even though there was no reference to them in the Delphi process, is a limitation in this study. A second limitation was the choice to use a convenience sample of OTs and PTs to interview to generate these codes. The involvement of a more diverse group of professionals similar to the Delphi Process would have been more ideal.

The importance of identifying general therapy behaviours has been evident in the development of generic fidelity measures in the psychotherapy literature (Di Rezze et al., In Press). General therapy behaviour characteristics have been described in some paediatric rehabilitation literature (Coster, Tickle-Degnen, & Armenta, 1995), however, the behaviours described have been relevant to a specific intervention (i.e., sensory integration).

The general therapy attributes generated from this study are the first set of behaviours relevant across a broad scope of health professions and paediatric rehabilitation interventions. With the multidisciplinary nature of paediatric rehabilitation, these attributes can be used to develop generic fidelity measures to assess general therapy

behaviours within intervention studies. These attributes could vary in importance based on the type of intervention or the population of children for whom it is delivered. Currently, the evaluation of intervention fidelity emphasizes the importance of differentiating interventions but does not assess characteristics common to all interventions. By examining all characteristics of an intervention delivered, researchers can better understand the mechanism of therapeutic effects on outcomes based on principles of FCS. The development of measures to assess general therapy attributes has the potential to improve the assessment of intervention fidelity in paediatric rehabilitation.

Future Directions

Future research can use these attributes to develop and test items for an observational fidelity measure in paediatric rehabilitation. Following the development of such a measure, work can begin to explore the relationship of each item to outcomes and to test whether attributes differ in their impact on clinical outcomes. Since this work was developed in the context of children with physical disabilities, other research should be conducted to explore the relevance of these attributes to children within other childhood disability populations (e.g., autism spectrum disorders).

References

Abraham, C., & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychology, 27,* 379-387.

Adami, M., F. & Kiger, A., (2005). The use of triangulation for completeness purposes. *Nurse Researcher, 12,* 4, 19-29.

Albarracin, D., Gillette, J.C., Earl, A.N., Glasman, L.R., Durantini, M.R., & Ho, M.-H. (2005). A test of major assumptions about behavior change: A comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. *Psychological Bulletin, 131,* 856-897.

Breitenstein, S.M., Gross, D., Garvey, C.A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation Fidelity in Community-Based Interventions. *Research in Nursing & Health, 33,* 164-173.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, A., Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, 2:40, doi: 10.1186/1748-5908-2-40

Casey, D. & Murphy, K. (2009) Issues in using methodological triangulation in nursing research. *Nurse Researcher, 16,* 40-55.

Coster, W., Tickle-Degnen, L., & Armenta, L. (1995). Therapist-Child interaction during sensory integration treatment: Development and testing of a research tool. *Occupational Therapy Journal of Research, 15,* 17-35.

Dickens, K., Matthews, L.R., & Thompson, J. (2011). Parent and service providers' perceptions regarding the delivery of family-centred paediatric rehabilitation services in a children's hospital, *Child: Care, Health and Development,* 64–73.

Dillman, D.A. (2007). *Mail and internet surveys: the tailored design method*. Hoboken, NJ: John Wiley & Sons, Inc.

Di Rezze, B., Law, M., Gorter, J.W., Eva, K., and Pollock, N., (In Press). A narrative review of generic intervention fidelity measures. *Physical and Occupational Therapy in Pediatrics.*

Dunn, W. (2011). *Best practice occupational therapy for children and families in community settings.* Philadelphia: Slack Books Inc.

Dunst, C.J. (2009). Implications of Evidence-Based Practices for Personnel Preparation Development in Early Childhood Intervention. *Infants and Young Children, 22,* 44-53.

Dunst, C.J., Trivette, C.M., & Hamby, D.W. (2007). Meta-analysis of family-centred help giving practices research. *Mental retardation and developmental disabilities research reviews, 13,* 370-378.

Elo, S. & Kyngas, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62,* 107-115.

Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28,* 1167-1187.

Ezzy, D. (2002). *Qualitative Analysis: Practice and Innovation.* London: Routledge.

Faw, L., Hogue, A., and Liddle, H.A. (2005). Multi-dimensional implementation evaluation of a residential treatment program for adolescent substance abuse. *American Journal of Evaluation, 26,* 77–94.

Foster, M., Whitehead, L., & Maybee, P. (2010). Parents' and health professionals' perceptions of family centred care for children in hospital, in developed and developing countries: A review of the literature. *International Journal of Nursing Studies, 47,* 1184-1193.

Graneheim, U.H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today, 24,* 105-112.

Green, J. (2006). Annotation: The therapeutic alliance – a significant but neglected variable in child mental health treatment studies. *Journal of Child Psychology and Psychiatry, 47,* 425 - 435.

Hong, C.S., & Palmer, K. (2003). Occupational therapy and physiotherapy for children with disabilities. *Journal of Family Health Care, 13,* 38-40.

Hsieh, H., Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15,* 1277-1288.

Jones, J.M., & Hunter, D. (1995). Consensus methods for medical and health services research. *BMJ, 31,* 376-80.

Junious D, Malecha A, Tart K, Young A (2010). Stress and perceived faculty support among foreign-born baccalaureate nursing students. *Journal of Nursing Education, 49,* 261-70.

King, G.A., King, S.M., and Rosenbaum, P.L. (1996). Interpersonal aspects of care-giving and client outcomes: a review of the literature. *Ambulatory Child Health, 2,* 151-160.

King, S., Rosenbaum, P., & King, G. (1995). The Measure of Processes of Care: A means to assess family-centred behaviours of health care providers. Hamilton, ON: McMaster University, Neurodevelopmental Clinical Research Unit

Law, M., Darrah, J., Pollock, N., Wilson, B., Russell, D., Walter, S., Rosenbaum, P., Galuppi, B. (2011). Focus on function: a cluster, randomized controlled trial comparing child- versus context-focused intervention for young children with cerebral palsy. *Developmental Medicine and Child Neurology,* 53, 621-629.

Law, M., Hanna, S., King, G., Hurley, P., King, S., Kertoy, M. & Rosenbaum, P. *(*2003*)* Factors affecting family-centred service delivery for children with disabilities. *Child: care, health and development, 29, 357–366.*

Lynn, M.R. (1986). Determination and quantification of content validity. *Nursing Research, 35,* 382-85.

McKenna H.P. (1994). The Delphi technique: a worthwhile approach for nursing? *Journal of Advanced Nursing, 19,* 1221–1225.

Morse, J. M. & Richards, L. (2002). *Read me First for a User's Guide to Qualitative Methods.* New Delhi: Sage

Nelson, D.L., & Mathiowetz, V. (2004). Randomized controlled trials to Investigate occupational therapy research questions. *The American Journal of Occupational Therapy, 58,* 24-34.

Rosenbaum, P., King, S., Law, M., King, G., and Evans, J. (1998). Family-Centred Service: A conceptual framework and research review. *Physical & Occupational Therapy in Pediatrics, 18,* 1, 1-20.

Sandelowski, M. (2000). Whatever happened to qualitative description? *Research in Nursing & Health, 23,* 334-340.

Verma, S., Paterson, M., & Medves, J. (2006). Core competencies for health care professionals: what medicine, nursing, occupational therapy, and physiotherapy share. *Journal of Allied Health, 35,* 109-115.

Whyte, J. & Hart, T. (2003). It's more than a black box; It's a Russian doll: Defining rehabilitation treatments. *American Journal of Physical Medicine and Rehabilitation*, 82, 639-652.

Table 1. Frequency count of categorized attributes after each round of Delphi Process

Round	Attributes rated	Attributes kept (consensus)	Attributes eliminated (consensus)	Attributes eliminated due to no consensus after two rounds	Attributes to be rerated (no consensus)
Round 1	35	24	0	0	11
Round 2	37	27	1	5	4
Round 3	36	35	0	1	0

Table 2. Final list of attributes from the Delphi Process

THERAPIST BEHAVIOURS
*1. THERAPIST assesses changes for the CHILD (formally or informally), e.g., child's functioning in daily life or skills taught in intervention.
*2. THERAPIST compliments and/or praises the CHILD'S effort or performance within the intervention session.
3. THERAPIST communicates confidence and/or acknowledges the CHILD'S gains during the intervention session
4. THERAPIST inquires about or discusses the availability and nature of support with the family, including supports available to the CHILD and PARENTS from outside the family
*5. THERAPIST discusses, reviews, or reformulates the CHILD'S (and/or family's) goals for therapy.
*6. THERAPIST actively engages the CHILD (and/or PARENT) in working together during the intervention session, including collaboratively formulating and following a specific plan for the session.
*7. THERAPIST develops and/or reviews one or more specific assignments for the client to carry out between intervention sessions
*8. THERAPIST conveys that he/she understands the CHILD'S problems, provides hope and/or demonstrates that he/she is able to help the CHILD within the intervention session.
9. THERAPIST conveys warmth towards the CHILD within the intervention session.
10. THERAPIST shows empathy toward the CHILD (and/or PARENT) within the intervention session.
11. THERAPIST teaches the CHILD skills or strategies during the intervention session.
*12. THERAPIST encourages the CHILD to engage in activities which would be pleasurable (e.g., toys, fun activities), and are appropriately graded for the CHILD in the intervention session.

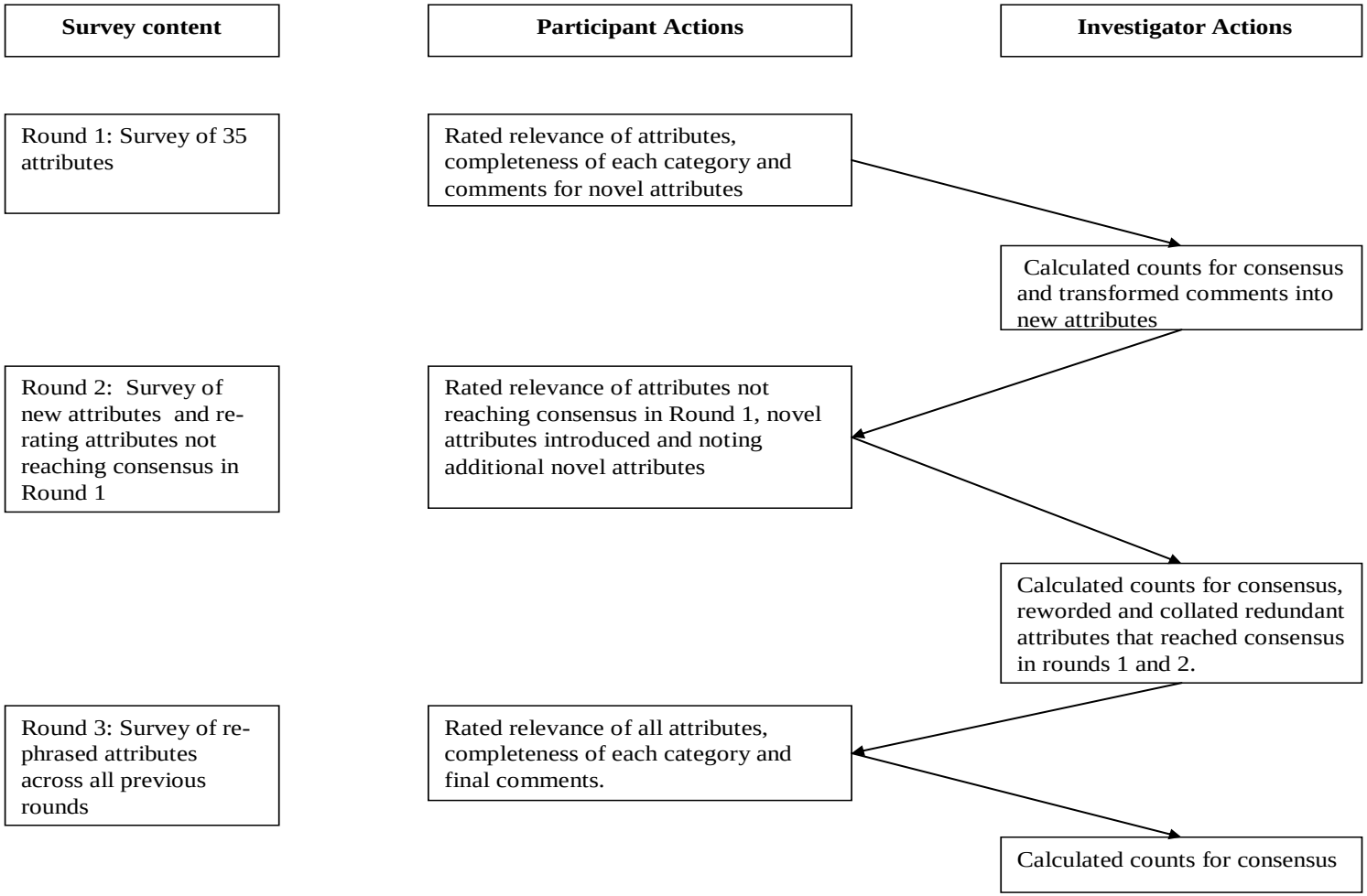
13. THERAPIST attempts to establish and/or maintain the CHILD's interest in the intervention within session.
14. THERAPIST presents knowledge and/or expectations about child development appropriate for the child within the intervention session.
15. THERAPIST focuses on exploring changes with the PARENT within the intervention session.
*16. THERAPIST listens, explores and/or supports the issues or information needs raised by the CHILD and/or PARENT within the intervention session.
*17. THERAPIST alters his/her behaviour to adapt to the challenges, needs, behaviours, interests of the CHILD in the intervention session.
*18. THERAPIST offers the CHILD a choice of activities within the intervention session.
19. THERAPIST shows appropriate closure to the intervention session.
20. THERAPIST models intervention and coaches parents about intervention and relevant issues.
*21. THERAPIST communicates in a way that is clear, professional and appropriate for the CHILD and PARENT in the intervention session.
CLIENT BEHAVIOURS
22. CHILD is receptive to participating or engaging in the intervention session.
23. CHILD shows confidence in the therapist or comfort with the intervention within the session.
24. CHILD communicates his/her preference(s) in the intervention session.
25. PARENT communicates his/her preference(s) in the intervention session.
26. CHILD or PARENT practice strategies with the therapist in intervention session.
27. PARENT is receptive to engaging or participating within the intervention session.
28. PARENT seems to show comfort within the intervention session.
29. CHILD shows that he/she was free of pain or discomfort during the intervention session.
*30. CHILD is actively engaged in the intervention session.
INTERACTION BETWEEN THERAPIST AND CLIENT
*31. THERAPIST and CHILD show a strong rapport in the intervention session.
32. THERAPIST and the CHILD reach consensus on decisions within the intervention session.
*33. THERAPIST and PARENT show mutually respectful behaviours in their discussions within the intervention session.
34. THERAPIST and CHILD actively communicate in order to work together within the intervention session.
*35. PARENT and THERAPIST come to consensus about how to proceed or what changes to make within the intervention session.

Note. * Indicates if the coding scheme mapped onto the attribute

Table 3. Sample of categories and codes used to analyze the Therapist Interviews

Category (bold) and codes	Definitions and Examples
1.0 Motivating the child	Therapist incorporates specific activities and/or games into therapy that capture the interest of the child.
1.1 Therapist provides activities that the child finds enjoyable.	Activities/games provided by the therapist enable the child to have a pleasant experience.
1.2 Therapist provides activities that intrigue the child.	Activities/games provided by the therapist are new to the child and captures his/her attention.
1.3 The session is fun for the child.	Overall, the child enjoys the session during and in between therapy activities
2.0 Therapist-child rapport	Child and therapist are able to get along well during an intervention session.
2.1 The therapist and child have a good relationship.	Child is responsive to the therapist and shows trust in therapist by engaging in activities within intervention session. E.g., child attempts new activities that therapist introduces.
3.0 Empowering the child	Demonstrates a respect for the child's autonomy (client-centred principle) to facilitate a comfortable working environment.
3.1 Therapist provides choices to the child, considers the child's preferences and/or negotiates with them as necessary.	Therapist enables the child to be independent in decision making within the session and discusses any limitations of options in order to accomplish intervention session goals.
3.2 Therapist follows child's lead during therapy session	Child able to do activities on their own (to whatever extent possible) or is encouraged to actively participate in session.

Figure 1. Flowchart of participant and investigator procedures for each round of Delphi Process



Chapter Four

Development of a generic fidelity measure for rehabilitation intervention research for children with physical disabilities

Authors: Briano Di Rezze, Mary Law, Kevin Eva, Nancy Pollock, Jan Willam Gorter

This chapter contains a manuscript entitled "Development of a generic fidelity measure for rehabilitation intervention research for children with physical disabilities". This manuscript has yet to be submitted to the proposed journal titled, Developmental Medicine and Child Neurology (DMCN). Its format in this chapter is in its expanded form, since it is not yet in the format for the requirements of this journal.

Abstract

Background: To advance paediatric rehabilitation research, there is a need to increase the rigour of intervention studies through clearly defining and measuring the delivery of an intervention. Investigators are beginning to measure intervention fidelity, defined as the degree to which interventions are delivered as intended. Fidelity measures developed in paediatric rehabilitation examine the active ingredients of a specific intervention. Generic fidelity measures are characterized by the evaluation of the active ingredients of more than one intervention, as well as general therapy characteristics (e.g., rapport) observed in all interventions. This study describes the development of the Paediatric Rehabilitation Observational measure of Fidelity (PROF), the first generic fidelity measure in paediatric rehabilitation for interventions involving children with cerebral palsy (CP).

Methods: The PROF was based on the Implementation Fidelity Framework. Items were constructed from content valid sources for a general attributes domain of paediatric rehabilitation and domains for two interventions. The intervention domains were applicable to children with CP (1-6 years) by an occupational therapist (OT) or physiotherapist (PT). From an RCT examining two therapies(one being a child-focused therapy aimed at changing the child's ability, the other being a context-focused therapy involved modifying the task and/or child's environment), over 50 videotaped intervention sessions were available for raters. Following an extensive pre-testing procedure, raters were trained and entire videos were viewed over the course of 4 weeks to test the measure's psychometric properties.

Results: The PROF consisted of 30 items – General attributes (20), Child therapy (4), Context therapy (4), and Global items (2). Most items were rated on two scales: Frequency and Quality of performance. Six raters (4 OT and 2 PT) were trained on the PROF and rated over 25 videos. Across the 6 raters, reliability results (internal consistency and inter-rater reliability) for the frequency scale of each domain showed: $\alpha = .71$, $G = .75$ (Context therapy); $\alpha = .85$, $G = .87$ (Child therapy); and $\alpha = .78$, $G = .82$ (General Attributes). Quality scale scores across domains demonstrated $\alpha > .80$ and IRR $< .40$. Construct validity testing showed that each of the three domain constructs were independent based on Pearson correlations and ANOVA analyses ($p = .01$). Pearson correlations showed correlations of scale dimensions (frequency and quality) for the intervention domains ($r \leq .50$) and General attributes $r = .63$.

Conclusions: This study demonstrated that the PROF is a reliable and valid tool for children with CP and the described interventions. Future studies involving children with CP may utilize the PROF framework, general attributes domain and procedures to test the psychometrics of other intervention approaches to ease the burden of fidelity measurement.

Introduction

An emerging, important area of measurement in non-pharmacological intervention research (including rehabilitation) is the evaluation of fidelity to the intervention (Boutron, Moher, Altman, Schultz, & Ravaud, 2008; Hennessey & Rumrill, 2003). Intervention fidelity is the degree to which an intervention is delivered as intended (Carroll et al., 2007). Evaluating intervention fidelity in rehabilitation research can ensure an accurate examination of a studied intervention and replicable comparison of specific interventions across study therapists and settings.

In the literature, intervention fidelity has been examined through the use of indirect (e.g., self-report), and/or direct methods (e.g., behavioural scores from observations) (Schlosser, 2002). Direct methods of examining intervention fidelity provide more accuracy in measurement and minimize altered perception of past performance, recall difficulties, and distortions in self-representation (Perepletchikova, Treat & Kazdin, 2007).

Generic fidelity measures directly evaluate both intervention-specific characteristics as well as general attributes of interventions (i.e., non-specific) that could moderate the intervention outcome. Intervention-specific characteristics examine the extent to which practitioners implement the active ingredients of the intervention(s) (i.e., prescribed attributes), as well as attributes that are excluded from the intervention (i.e., proscribed attributes). General attributes of intervention are characteristics considered essential to all interventions (e.g., therapist rapport with a client) (Breitenstein et al., 2010). General attributes should be evident during all intervention sessions and are not

specific to one intervention approach. Knowledge of the presence of intervention specific and non-specific behaviours can enable an investigator to examine whether there is an interdependent relationship between them, and to what extent the general therapy behaviours impact on client outcomes.

Generic fidelity measures demonstrate traits that can decrease the burden on individual investigators. A generic measure would enable generalizability across different interventions that are consistent in their conceptual foundations (Breitenstein et al., 2010). Generic measures can also reduce the cost and time resources of developing new fidelity measures since they provide a foundation or framework to be applied to novel interventions.

A narrative review of generic fidelity measures by Di Rezze, Law, Gorter, Eva and Pollock (In Press) used in adult psychotherapy found that therapist and client behaviours as well as their interaction are general attributes evident in any intervention session. These general attributes are important to examine for intervention fidelity measurement, since they are reported in the literature as potential moderators of intervention (Carroll et al., 2007; Web, DeRubeis & Barber, 2010).

In the paediatric rehabilitation literature, a generic measure of fidelity, including general and specific attributes, has yet to be developed. This study describes the development and initial psychometric testing of a generic fidelity measure of rehabilitation intervention sessions for children with cerebral palsy.

Methods

Several methods were used to ensure that the fidelity measure was grounded in theory and systematically constructed following the recommended steps to develop a fidelity instrument (Stein, Sargent, & Nicholas, 2007). These methods included: (1) definition of the foundational constructs of the measure; (2) development of the measure (i.e., item development, scaling and initial testing); (3) establishment of scoring procedures; and (4) examination of reliability and validity. The development of this measure occurred within the context of a large randomized controlled trial (RCT) comparing two rehabilitation interventions for children with cerebral palsy (briefly described in the next section) (Law et al., 2007; Law et al., 2011). Ethics approval was obtained for all parts of this study through the Faculty of Health Sciences Research Ethics Board at McMaster University.

1) Definition of the Foundational Constructs of the Measure

a) Construct of Fidelity

Coinciding with the principles of generic fidelity measures that examine prescribed, proscribed and non-specific behaviours, the generic fidelity measure developed for this study was also based on a conceptual model - the Implementation Fidelity Framework (IFF) (See Figure 1) (Carroll et al., 2007). In this model, therapist adherence in delivering the intervention mediates intervention fidelity. The IFF also identifies potential moderators of fidelity that specify for whom and under what conditions a higher degree of fidelity is achieved.

(Kraemer et al., 2002; Baron & Kenny, 1986). Potential moderators (listed in Figure 1) may occur prior to the delivery of the intervention in a study, such as considering the comprehensiveness of policy (or the complexity of an intervention) and defining therapy strategies to better facilitate implementation. Other moderators can be examined within an intervention session, such as the quality of delivery and participant responsiveness. Quality of delivery refers to the proficiency of the therapist to deliver the intended intervention, whereas participant responsiveness includes factors related to both the deliverer (i.e., therapist) and recipient (i.e., client) (Carroll et al., 2007). The fidelity measure developed in this study focused on the evaluation of both intervention-specific attributes and general attributes within an intervention session. Specifically, the fidelity measure includes a domain for each of the two compared interventions and one domain for the general attributes of an intervention session for children with cerebral palsy.

[Insert Figure 1 about here]

b) Constructs of Intervention-Specific Domains

The constructs defining the two intervention-specific domains in the measure were derived from a large multi-site RCT called the Focus on Function study (FOF) (Law et al., 2007). This study compared the efficacy of two approaches to paediatric rehabilitation intervention to improve performance of functional tasks and mobility and increase participation in everyday activities for children with cerebral palsy (12 months to 5 years of age) from 21 children's treatment centres across two provinces of Canada. A full description of the study and results are available from Law et al. (2011) and Darrah et al. (2011). Within the FOF study, occupational therapists (OT) and physiotherapists (PT)

provided intervention using one of two approaches: (i) Child-focused therapy (or Child therapy) - remediation of impairments and improving the child's skills and abilities; and (ii) Context-focused therapy (or Context therapy) - changing constraints within the task or environment to improve performance.

Each intervention approach included unique strategies that were the "prescribed" attributes for the interventions. Intervention strategies not included in each approach are the attributes that are "proscribed" or excluded. The examination of both prescribed and proscribed attributes ensures the examination of the overall integrity of the delivered intervention. Therapists conducting the intervention in the FOF study (n=79) were asked to videotape a "typical" intervention session at about the midpoint of their 6 month intervention block. This video was examined by FOF study investigators to ensure that they conducted the intervention as trained. Using videotaped sessions is a common and accepted method in intervention fidelity studies (Mowbray, Holter, Teague, & Bybee, 2003). From the study, 67 therapists provided a videotaped intervention session. Of this pool of videos, only 39 were usable for this study based on a priori exclusion criteria. Videos were not used for this study if: they had video quality issues (n=8); were videotapes of goal development sessions rather than intervention (n=17); or if the therapist was not directly delivering the intervention on video (n=3).

For the intervention-specific domains of Child or Context focused therapies, attributes were developed by reviewing descriptive literature on the interventions (Darrah, Law and Pollock, 2001; Law et al., 1998; Law et al., 2007) and therapist training materials from the FOF study. This review was used to identify the essential attributes of

each intervention and how the interventions were distinct, following a process used successfully in the development of another fidelity measure (Carroll et al., 1998).

c) Constructs of the General Attributes Domain

The constructs underlying the general attributes of a paediatric rehabilitation intervention session were established by two studies that identified 3 therapeutic behavior categories (Di Rezze et al., In Press) and the key characteristic behaviours for each (Di Rezze, Law, Eva, Pollock, & Gorter, Submitted). The first category of behaviours considered essential for any intervention for children with physical disabilities is therapist behaviours, including content and procedure related behaviours. The second category describes behaviours of the client, primarily characteristics of the child's or parents' responsiveness and participation within the intervention. The third category involves client-therapist interaction behaviours including interactions between the therapist and the child or parent. The findings of the Delphi study identified 35 general attributes that were considered essential for generic fidelity.

2) Development of the fidelity measure

a) Construction of items

During item generation, the attributes were converted to observable items for each of the domains: (1) Child Therapy; (2) Context Therapy; and (3) General Attributes of Therapy. Intervention-specific items were initially drafted by the primary investigator (BD) and combined with descriptive examples from FOF study materials. Further refinement of these items took place through observation of sample videotapes identified by the FOF investigators as ideal examples of each intervention. This method has been

used by other fidelity measure developers to refine items of observed behaviours on videotaped (or audio taped) sessions (Waltz, Addis, Koerner, & Jacobson, 1993). To develop items in the General Attribute domain, the principal investigator modified the wording of the attributes from the Delphi study to ensure that statements were observable. Examples of types of therapist, child or parent behaviours considered an "observable behavior" were added to clarify the meaning of the attribute(s).

Following the item drafting process, consultations with the 6 FOF investigators verified the face validity of the items specific to each therapy domain. Final refinements of the items ensured that item language was consistent across all domains, and that items could be observable in a videotaped intervention session. A rater's manual was developed following a format similar to three prominent generic fidelity measures by Martino, Ball, Nich, Frankforter, & Carroll, 2008; Hollon, 1984; and Diamond, Diamond, & Hogue, 2007. Rater instructions detailed definitions of each domain with examples, descriptions of the scaling dimensions, and procedures for scoring.

b) Item Scaling

Common dimensions in generic fidelity measures include "adherence of performance" scale to assess presence and frequency of the behaviour, and a "competence" scale to assess quality of delivery behaviours (Carroll et al., 2000; Martino et al., 2008). Both dimensions were used to rate each item in the fidelity measure. In existing measures, adherence scales have been defined differently, from identifying the frequency of behaviour (Godfrey, Chalder, Ridsdale, Seed , & Ogden, 2007) to combining frequency with intensity of behaviour (Hogue, Liddle, Singer & Leckrone,

2005). For this fidelity measure, adherence focuses on frequency of behaviour. Quality scales have typically been used uniformly across measures, focusing on how well a behaviour has been conducted (Carroll et al., 2000).

The selection of scaling for each dimension was based on examining fidelity as a degree of frequency or quality, rather than a dichotomous concept (Breitenstein et al., 2010; Carroll et al., 2007). A continuous direct estimation approach was used with a 5-point scale using adjectives consistent with observed performance (see Figure 2). The use of 5-points on each scale respects the limits of human discrimination across dimensions (7 +/- 2 boxes) from the concept of Miller's Magic Number (Streiner & Norman, 2006). Scores move from 1-5 in adherence scoring and 1-5 for quality. An item is scored "N/A" if it is not observed. Items that are scored 1 or N/A for adherence are also scored N/A on quality. Similar scaling for each dimension (odd numbered and continuous intervals) has been used successfully to rate each item within other generic fidelity measures (Carroll et al., 2000; Martino et al., 2008).

[Insert Figure 2 about here]

c) Pretesting the Fidelity Measure

The stages of testing the measure are outlined in Figure 3. In Stage 1, the measure and its manual were examined by a subsample of investigators from the FOF study and clinicians familiar with the study interventions. Three group meetings were conducted involving: two paediatric therapists (providing input on item clarity); two FOF study investigators with experience in developing observational measures (commenting on item definitions and descriptions); and three FOF study investigators who developed the

interventions (evaluating intervention items and their distinctions). Within each meeting, discussion of areas for which there was disagreement was led by investigators BD and ML to achieve consensus. Participants in this initial testing phase were introduced to the measure and manual, rated two videotapes (one from each intervention), and provided feedback on item clarity, item structure and utility (Barber & Crits-Christoph, 1996).

[Insert Figure 3 about here]

d) Description of the Paediatric Rehabilitation Observational measure of Fidelity (PROF)

The generic fidelity measure included 3 domains (i.e., two therapy domains and one general attributes domain) for a total of 28 items. In addition, two global scales were included to examine overall ratings of fidelity and quality. Each therapy domain had 4 items and the General Attributes domain included 20 items. The General Attributes domain included general therapist behaviours (8 items), child behaviours (7 items), and behaviours of the parent (5 items). All 28 items were rated on a "frequency of behaviours" scale, and all items examining solely therapist performance (n=18 items) were also rated on the "quality of delivery" scale. See Appendix B for the complete measure and rating manual.

3) Establishment of Administration Procedures

a) Rater Criteria and Training

Due to the complexity of the task, raters were required to be experienced paediatric therapists (OTs or PTs) (Moras & Hill, 1991). Eligibility criteria for therapist experience consisted of more than five years of experience (King et al., 2008) conducting

interventions involving children with physical disabilities. Written consent was collected before they participated in the study. Videotaped intervention sessions were evaluated by 6 raters independent of the FOF study in order to minimize social desirability bias (Mowbray et al., 2003). Raters were also naive to the purpose and hypotheses of this study.

In the literature, rater training have typically occurred through didactic seminars and rating practice of least 10 videotapes which are compared to expert consensus (Carroll et al., 2000; Martino et al., 2008). The training for the PROF consisted of approximately 16 hours involving didactic instruction and additional ratings of videotaped interventions (10 videos in total), see Stage 2, Figure 3.Within this training period, participants rated videos independently and feedback was collected weekly to identify any difficulties. A summary of all rater queries were sent weekly to all raters.

Over the three weeks, videos were scored until each rater reached acceptable agreement (75%) in terms of proportion of codes rated correctly for each intervention type. Agreement was defined as within two scale points in comparison to the standard score of expert consensus (Martino et al., 2008) set by the average group scores. Remediation was provided to a rater if agreement was not reached. Remediation consisted of a telephone meeting to discuss rater rationale for their responses and use of the manual as a guide to recalibrate participant ratings.

As part of feasibility testing, rater feedback was used to improve item and scaling clarity. Item ratings for these videos were used to conduct preliminary analyses to

examine endorsement frequencies of items, exploratory factor analysis of item domains, and early reliability data.

b) Scoring procedure

The observational unit for rating each video was based on the "occurrence and non-occurrence approach" (Waltz et al., 1993, p.623), whereby the rater reviewed the session in its entirety to determine whether a behaviour occurred, how often it occurred and the quality of its occurrence. Full video review for scoring ensured that crucial information was not missed as this can occur if video segments are used (Gardner, 2000). The average score of each intervention specific domain and the General Attribute domain within the fidelity measure indicates the degree of fidelity of each domain. Higher scores would indicate better fidelity.

For the intervention specific domains, a scoring procedure was developed to score prescribed attributes that were expected as well as proscribed attributes that were not expected. For example, in a Child focused therapy video, a therapy activity to change the child's skills is a prescribed attribute while an activity involving the modification of the environment is proscribed. A high fidelity Child focused therapy video would include therapist behaviours rated higher in the Child therapy domain, and lower in the Context therapy domain. To ensure that both prescribed and proscribed items were scored appropriately, a separate analysis was conducted by combining both therapies and using reverse scoring for items in the domain opposite to the video being rated. As a result, scaling for domain items consistent with the video type were scored in increasing value (1 to 5), and decreasing value for items inconsistent with the video (5 to 1).

Reverse scoring was not applied to item ratings in the General Attributes domain, because these behaviours were expected across both interventions. The global fidelity score was a 5-point scale that asked participants to rate the overall purpose of the therapy observed on video from 1 = 'Change the environment' (implying context therapy) to 5= Remediate impairments (implying child therapy).

4) Examination of Reliability and Validity

For reliability and validity testing, raters scored 25 videotaped intervention sessions over 4 weeks (Stage 3, Figure 3). These ratings were used to test the reliability (i.e., Inter-rater reliability and internal consistency) and validity (i.e., convergent and discriminant) of the measure.

To calculate the number of videos to be scored by each rater, a sample size estimation formula was used that required a minimally acceptable Intraclass Correlation Coefficient (ICC), and the number of desired raters (Stratford & Spadoni, 2003). Based on the reliability coefficients of other fidelity scales, r = 0.71-0.85 (Carroll et al., 2000; Eames et al., 2008), it was anticipated that a reliability of .85 was attainable and a minimal acceptable reliability would be r=.70. With an expected reliability of .85 and using six raters, approximately 11 videotapes for each intervention approach were expected to be required to be scored by each rater to achieve a 1-sided lower 95% confidence limit of .70 (Stratford & Spadoni, 2003).

To ensure that raters maintained a standard level of scoring following their training (i.e., quality assurance post-training), raters also scored a common video on two random occasions (two weeks apart) to monitor rater drift (Martino et al., 2008). For this

evaluation, individual rater domain scores on the second occasion were compared to the average group domain scores from the first rating occasion. If drift was observed, remediation was done by telephone to highlight inconsistent responses and provide feedback on how to respond to items more consistently. Over the course of psychometric testing, a counterbalancing method was used to ensure raters scored different videos each week, and the order of four sets of videos was varied.

Data Analysis

Demographic information about raters' clinical experience including, OT/PT practice setting, and client characteristics, were described using descriptive statistics. Similar statistics were used to describe demographic data for the videotaped therapist (i.e., profession type and intervention delivered) and child (i.e., age and level of gross motor ability). Descriptive data and qualitative comments about the measure were also collected, such as time to rate each video and experiences in using the PROF. Endorsement counts were considered to examine the occurrence and dispersion of scores across items. Quality assurance data using means and 95% Confidence Intervals were collected to evaluate rater consistency at two time points within the study. Standard drift in the fidelity literature was indicated by scores outside of the 95% confidence interval for the rater group (i.e., more than two standard deviations from the group means for each domain) (Martino et al., 2008). A more sophisticated data analysis examined rater drift over the course of the four rounds of ratings using Analysis of Variance (ANOVA).

ANOVAs were also conducted to examine whether rater scores differed by rater profession.

Generalizability (G) theory was the conceptual/statistical approach used to analyze the reliability of the data (Streiner & Norman, 2006). Reliability between items (internal consistency) and between raters (Inter-rater reliability) was calculated separately for the three domains in the measure using video as the facet of differentiation. For the analyses, the main source of expected error was differences in rater scoring. Inter-rater reliability calculations were interpreted based on the following standards: Poor $\leq$.40; fair = .40 to .59; good =.60 to .74; and excellent $\geq$.75 (Cicchetti, 1994).

For inter-rater reliability, G-coefficients were generated for the frequency and quality of delivery scales in each domain, also global items (fidelity and quality). Other reliability analyses were conducted by examining data by intervention-type (on video) and rater profession. A Decision Study (D-Study) was conducted if G-coefficients did not meet the .70 minimum or exceeded .90 to determine the number of raters and/or items needed to achieve this standard in future studies.

Measuring the internal consistency of instruments in intervention fidelity studies compares the relationships between items within each domain (Bond et al., 2002). Cronbach's alpha (α) coefficient for each therapy domain was expected to be between .70-.90 (Streiner & Norman, 2006). Items in the General Attributes domain were not expected to be homogeneous, since they represent different attributes; intervention-specific items within each domain were expected to correlate with each other between r=.20-.80 (Streiner & Norman, 2006).

To test the validity of the fidelity measure, two sets of construct validity tests were conducted based on a priori hypotheses. The first set of analyses examined discriminant validity between the dimensional scales - frequency and quality of delivery, using Pearson correlation coefficients. The second set examined the validity of each domain construct and the expected relationships (i.e., converging or discriminating) between each domain. This analysis involved calculating Pearson correlation coefficients and ANOVAs across videos and parceling videos out by intervention type. Effect sizes were calculated for ANOVAs, using eta scores and the following categories: .01 = small: .06 = moderate; and .14 = large (Norman & Streiner 2008, p.87). Additional Pearson correlations were conducted using video ratings with a separate data source (i.e., clinical chart notes by the clinician on video) from the FOF study. For this analysis of convergent validity, data from video ratings were compared to frequency data from coded clinical session notes written by the treating therapist on the same dates as the videos (Law et al., 2011 or Darrah et al., 2011). Session notes were coded by an independent rater with experience in coding clinical notes from the FOF study.

Preliminary Exploratory Factor Analyses (EFA) were conducted to examine how well domain items loaded together and differently from other domains. These results were considered preliminary because of a lack of sufficient data points to conduct EFA.

Results

Demographics

A total of 6 raters were trained to conduct the psychometric testing of 25 intervention videos (14 Child therapy and 11 Context therapy interventions). The average

amount of experience for the group of raters was 14 (SD=6) years, working in a wide range of settings (i.e., children's treatment centres to community contexts), and the client age range was 0 to 21 years. Four raters were OTs and two were PTs. Raters reported that they primarily worked with children with physical disabilities. Common diagnoses seen by these practitioners were cerebral palsy, spina bifida, developmental delay and Down's Syndrome.

Descriptive information about the participants and therapy on video was summarized. Across the 25 videos the child mean (SD) age was 3.9 (1.3) years with a range of 1.5-6.2 years. The motor abilities of children on video were identified using the Gross Motor Function Classification System (GMFCS) (Palisano et al., 1997). Children were grouped into either GMFCS levels 1 to 3 or levels 4 and 5. As a result, 60% (n=15) of children were labeled within levels 1 to 3, and 40% (n=10) possessed abilities of either level 4 or 5. In terms of the intervention on video, 56% (n=14) were child-focused interventions and 44% (n=11) were context-focused. Parents were present on 40% (n=10) of intervention videos. The professionals delivering the intervention on video consisted of 56% (n=14) OTs and 44% (n=11) PTs.

PROF Utility and Item Endorsement

Rater comments regarding their experiences in using the PROF were collected throughout psychometric testing. Specific comments indicated that raters felt more comfortable with the PROF as they rated more videos because they were more aware of what to look for on video. Collectively, the raters averaged 17 minutes to rate one video,

with their rating time slightly decreasing over the study from 20 to 15 minutes. The length of time for each videotaped intervention session ranged from 15 to 50 minutes. Raters commented on the variability of the video quality on some videos, citing specifically that low speaking volume made some difficult to follow. Raters also noted difficulties in rating some videos based on uncertainty of the goals addressed, and/or if the session focused more on examining a child's goal status versus an intervention session working on specific goals.

Endorsement counts across items scored on the frequency scale showed that across raters, almost all items were rated for each video. Items from the therapy domains demonstrated patterns of scores aligned with their respective intervention video across the 149 observations. Child therapy items were more often rated as "1" (behaviour not observed on video) in Context focused therapy videos (endorsement counts 6 to 15) in comparison to Child focused therapy videos (counts of 0 to 3). For Context therapy items, ratings of "1" were predominantly in Child focused therapy videos (counts of 19 to 41) compared to Context focused therapy videos (counts of 8 to 21). Most General Attribute items showed a similar occurrence and dispersion of scores across both types of intervention videos. However, the five General Attribute items describing parent behaviours showed patterns of "not observed" scoring more often in Child focused therapy videos (counts of 5 to 33), whereas in Context focused therapy videos, parents were most often in attendance (counts of 0 to 6).

For both types of intervention videos, two General Attribute items had high frequency scores of behaviours "not observed" on video: Items #10 and #20. For item

#10, "Did the therapist **review therapy goals and/or activities** for the CHILD (and/or Parent) **to practice outside of the intervention** sessions?" 52% (77/149) of the time, it was not observed. For item #22, "Did the CHILD **show discomfort due to a personal factor** (e.g., physical pain, sleepiness, emotional lability)?", 66% (98/149) of the time, it was not observed.

Upon examining rater drift, the mean and 95% Confidence Interval (C.I.) of the ratings assigned on the first video viewing were: 4(2.92,5.08) for Child focused domain; 2.29 (0.81,3.77) for Context focused domain; and 3.37 (2.35, 4.39) for the General Attributes Domain. Domain scores for each rater for the second viewing of the same video resulted in scores that fell within the CI for each respective domain for 5 raters. Rater 3 in Table 1 demonstrated drift for the Child focused therapy video, for which remediation was provided.

[Insert Table 1 approximately here]

To examine if variables such as rater profession, domain-type and/or scoring differences over time (over 4 rounds of rating) due to proficiency bias, a three-way ANOVA was used to test significant differences of mean domain scores. Independent variables for the three-way ANOVA were time (i.e., 4 rating rounds), domain (3 domains) and rater profession (OT vs. PT). The main effects for Round were $F = .47$, $p = .71$, for Domain were $F = 43$, $p < .001$ (eta = .23), and for Rater Profession were $F = .15$, $p < .70$. Domain score differences were not unexpected, since each domain measured different behaviours. The interaction effect of Round by Domain was $F = 1.8$, $p = .10$, Round by Rater Profession $F = .15$, $p = .93$, and Rater Profession by Domain $F = .88$, $p =$

.42. Means, standard error and C.I.s by round and rater profession for each domain are seen in Tables 2 and 3.

[Insert Table 2 approximately here]

[Insert Table 3 approximately here]

Reliability

Internal Consistency

For the internal consistency of the frequency scores for child therapy and context therapy domains, alpha coefficients were greater than .70 (Table 4). Both therapies collapsed together and reversed scored resulted in $\alpha = .73$. The α-coefficients of the quality of delivery scores for the therapy domains were greater than .80. The general attributes domain for the frequency scale had an α-coefficient of .78, while the quality scale ratings of the General Attributes domain was $\alpha = .81$.

[Insert Table 4 approximately here]

Using all videos (n=25), and **averaging across all 6 raters**, the inter-rater reliability of each domain for frequency scores were G-coefficients $\geq .75$. Reliability for frequency scores of all domains resulted in G-coefficients ranging from .75 for Context-focused therapy to .87 for Child-focused therapy (Table 5). For the reliability of a **single rating** (i.e., if only scored by 1 rater), G-coefficients ranged from .33 for context-focused therapy to .54 for child-focused therapy. Each domain for the frequency scale had complete data (n=149). When combining the therapy domains and reverse scoring based on intervention-type, the reliability analysis resulted in G = .80.

Additional analyses were conducted to examine the reliability of scores within each intervention type (Table 5). Inter-rater reliability for Child focused therapy videos had the lowest G-coefficients in all domains for frequency (G = .40) and quality (G = .33) ratings. Reliability analyses were also conducted for raters grouped by profession and intervention type. OT (n=4) reliability scores resulted in G = .69 for child therapy and G= .56 for context therapy. PT (n=2) reliability scores resulted in G = .55 for child therapy and G = .03 for context therapy.

Quality of delivery scores showed G-coefficients of .41 for Inter-rater reliability averaged across 6 raters. For the analysis, quality scoring demonstrated missing data points for each domain thus resulting in a reduced number of available items included in the analysis from 149 ratings to 114 for Child therapy , 59 for Context therapy and 45 for General Attributes.

Reliability analyses for the Global items (Fidelity and Quality) were also conducted across all videos. For Inter-rater reliability, the global fidelity item, rating the overall fidelity of the video, was G = .34 across the 6 raters. For global quality scores, Inter-rater reliability was G = .80. The equations used to generate the data for the G-Studies can be observed in Appendix C.

[Insert Table 5 approximately here]

From the reliability analyses, the majority of error variance was attributable to rater differences and several D-Studies identified the optimal conditions for future use of the measure to ensure improved reliability scores. The minimum number of raters required to achieve G > .7 within each domain are: 5 raters for Context therapy; 3 raters

for General Attributes domain; and 2 raters for the Child therapy Domain. D-studies that identified optimal number of raters by profession-type were also calculated (See Appendix D for the equation used for this D-study). To achieve G > .70, for OT, 5 and 8 raters are required for child and context therapy. The D-study for PTs showed 5 raters necessary to achieve a standard of G > .70 for child therapy. However, the D-study for context therapy indicated that G > .70 could not be achieved with less than 30 raters (G= .29). The complete tabled results data for this D-Study can be observed in Appendix E.

Validity

Validity of scaling dimensions - Distinguishing between Frequency and Quality

Pearson's Correlation coefficients (r) were used to determine if the frequency and quality scores used for therapist domain items demonstrated low correlations as expected. Each domain had the following correlations: Child therapy (r=.61, p=.001); Context therapy (r=.56, p=.004); and General Attributes (r=.63, p<.001).

The relationship between global rating scales for fidelity and quality demonstrated r= - .01, p= .95. Correlations between Global scores of quality and average quality scores for each therapy and general attributes domain were: Child therapy domain, r=.84, p<.001; Context therapy domain, r=.77, p<.001; and General attributes domain, r=.80, p<.001. To examine the relationship of the frequency and quality items, a series of Exploratory Factor Analyses (EFA) were conducted (Appendix F). For the child therapy domain (Appendix F, Table 1) the top 3 factors accounted for 78% of the variance. Factor

1 loaded all quality items together. Factor 2 loaded all child frequency items together. Factor 3 loaded both the frequency and quality scores for one child item.

In the EFA for context therapy items (Appendix F, Table 2) the top 3 factors accounted for 79% of the variance. All quality items loaded onto factor 1, in addition to one frequency item. Factor 2 showed two frequency and two quality items loading together. For Factor 3, three of four frequency items loaded together.

For the General Attributes domain, the first five factors accounted for 75% of the variance (Appendix F, Table 3). For each of the factors, except for factor 2, the frequency and quality dimensions for each item loaded together – factor 1 (items 15-18), factor 3 (3 items), factor 4 (1 item), and factor 5 (1 item).

Validity of the domain constructs

i. Predictability of the PROF based on domain type for all videos:

a) Video rating data analyses

Pearson correlations were calculated to examine the association between domain frequency scores (table 6). The relationship between the therapy domains was negative, $r = -.71$, $p<.001$. The General attribute domain correlations with Child and Context focused therapies were .43 ($p=.032$) and -.07 ($p=.73$), respectively.

[Insert Table 6 approximately here]

When examining the child and context therapy items together in an EFA (3 factors accounting for 75% variance), factor 1 loaded all child therapy items, whereas context therapy items loaded separately over factors 2 and 3, (Appendix F, Table 4).

b) Chart note data analyses

Therapist clinical chart notes also reported the frequencies of therapy behaviours of each therapy domain for 23 of the 25 videos. Table 7 summarizes the frequency counts of therapist behaviours related to either therapy approach. Pearson's Correlation between Chart context therapy frequency counts compared those of child therapy resulted in r= -.64 (p=.01).

[Insert Table 7 approximately here]

c) Video and Chart note data analyses

Therapist clinical chart notes were used to examine the relationship between the frequencies of these data and video data for child and context domains (Table 8). Reverse scoring of video frequencies were compared to the proportion of context and child frequency behaviours identified within the clinical chart note. Correlations for the reverse scoring method applied to the PROF with the proportion of child therapy behaviours from the chart resulted in r=.63 (p=.001), and r= - .62 (p=.002) with the proportion of context therapy behaviours.

[Insert Table 8 approximately here]

ii. Predictability of the PROF based on video-type

a) Domain frequency ratings for each video type

A two-way repeated measures ANOVA examining domain and intervention type for frequency scores revealed significant main effects of intervention type (F=8.41, p=.004, eta=.06), Domain (F=75.94, p<.001, eta=.36); and interaction of Intervention type by Domain (F=101.26, p<.001, eta=.42.

To clarify the source of these effects, one-way ANOVAs were conducted for each domain (Table 9). Statistically significantly higher scores were observed for child therapy domain scores for Child focused therapy videos compared to the scores of context therapy videos, $F=108.26$, $p<.001$, eta $=.42$. Context therapy domains also demonstrated a statistically significant difference by intervention type with higher means for Context focused therapy videos in comparison to Child focused therapy videos, $F= 38.96$, $p<.001$, eta $=.21$. The General Attributes domain did not demonstrate statistically significant differences between intervention type, $F=.65$, $p=.42$.

[Insert Table 9 approximately here]

(b) Global fidelity ratings

Global fidelity scores comparing intervention types (illustrated with means and standard deviations) for Child 4.37(.69) and Context therapy 2.46(.93), showed statistically significant differences, $F=34.81$, $p<.001$. Global quality ratings were not statistically significantly different between intervention types Child therapy 3.59(.60) and Context therapy 3.36(.77), $F=.69$, $p=.41$.

Discussion

The PROF is the first observational assessment in paediatric rehabilitation to examine a construct of fidelity that includes the evaluation of general attributes of the therapy process. This emerging perspective to measure fidelity accounts for the important roles of the client and therapist within an intervention session. Through the General Attributes domain, this measure incorporates and evaluates the dynamic nature of

paediatric rehabilitation by examining the therapist, child and parent behaviours - those typically involved in its family-centred nature. Furthermore, the therapy domains were examined together with reverse scoring, to rate prescribed and proscribed behaviours for each therapy, a characteristic of fidelity that has not been identified previously in measures of fidelity in paediatric rehabilitation (Parham et al., 2011; Bayona, McDougall, Tucker, Nichols, & Mandich, 2006).

Reliability

The generic fidelity measure examined in this study has good to excellent reliability for frequency scores across each domain, both for internal consistency (α = .70 to .85) and inter-rater reliability (G = .75 to.87). For internal consistency, this indicates that each domain contains items that are consistently related to the same underlying constructs. By convention, an acceptable estimate for domain consistency is $\alpha \geq .70$. Existing intervention fidelity research demonstrated typical values of α = .72 (Bond et al., 2002) and α = .74 (Lucca, 2000). For Inter-rater reliability, the frequency scores for each domain for General Attributes (G=.82) has exceeded reliability scores of other measures in the literature ICC= .67 (Diamond et al., 2007) or .39 (McIntosh et al., 2005).

In the literature, other fidelity measures have reported reliability analyses for each domain separately by intervention-type (Martino et al., 2008; Carroll et al., 2000). This conservative analysis has likely been conducted to minimize the variance that could result when combining the items of the different interventions. In addition to conducting such analyses, this study used a novel approach to minimizing this variance by combining the therapy domains and reverse scoring the ratings. This combined score also enables the

examination of the fidelity of prescribed and proscribed behaviours for each therapy. Testing resulted in an excellent Inter-rater reliability score (G=.80). This score provides confirmatory data on the ability to integrate scores from each intervention onto one scale that will be useful in future work to generate meaningful fidelity scores for an intervention session (i.e., the degree of fidelity achieved).

By intervention type, it was evident that Child focused therapy videos had lower reliability scores. Raters may have had more difficulties consistently differentiating between performances in the Child therapy domain when rating only Child focused therapy videos. This was likely due to decreased variability of performance on Child therapy videos, because Child domain scores showed good reliability when all video data were examined together, including when videos were stratified by rater profession (OT vs. PT). On the other hand, the Context therapy domain possessed sufficient variability to differentiate videos with good reliability, except for scores rated by only PTs ($G < .1$). This result is likely due to the Context therapy being a novel approach to PT raters, an approach that OT clinicians are more familiar in seeing and doing within their practice. Child therapy focuses on remediating a child's abilities and is the more traditional therapy approach familiar to both PTs and OTs in paediatric intervention.

For quality scores, inter-rater reliability was poor across all domains. In the literature, quality domains have been more reliable even when reliability testing was done more conservatively for each domain by intervention type, e.g., .71 to .98 (Carroll et al., 2000). The difference between the reliability of quality scoring in Carroll et al. (2000) and this study could be due to the psychotherapy interventions (and non-specific

domains) being well established with clear and specific descriptions of quality standards that are easier to score (Carroll & Nuro, 2002). However, this study demonstrated that the Global quality item had excellent reliability G= .80, providing support for the future use of overall judgments of quality for a video, instead of individual item ratings.

Validity

Findings from this study indicated that the PROF is a measure that possesses distinct constructs for each of its three domains, two comparator interventions and general attributes of intervention common to both interventions. Similarly, global ratings (fidelity and quality) were distinct from each other. Therapy domains were negatively correlated and statistically significant differences were evident between intervention-type for each domain. This finding supports the fidelity of the interventions (child and context therapies) that were distinct and strongly opposed to each other.

For General Attribute ratings, there were no significant differences between the intervention videos. This finding indicates that these items were common to both interventions and applied equally well across interventions. The correlations of therapy domains frequency scores with the General Attribute domain were weak with child therapy (r = .43) and very low for context therapy (r = -.07) domains. Although general attributes were evident in both intervention groups, Child therapy intervention had a stronger relationship with such items. The way in which Child therapy is delivered may enable the therapist to demonstrate more behaviour of general attributes (i.e., more direct therapy with the child). Alternatively, Context therapy was a novel intervention for the

therapists so it may have been more challenging for therapists to comfortably deliver the intervention and demonstrate it on a videotape.

Quality ratings for General Attributes domain moderately correlated with frequency scores ($r = .63$, $p = .01$). This finding indicates that quality ratings for this domain may not be importantly different from frequency ratings. The general attributes domain could implicitly possess characteristics of quality if it was observed on video. For example, if a therapist provides a higher frequency of praise to a child, higher quality of therapy could be assumed. This means that at an item level the quality ratings do not provide any helpful information. Hence, for future use of the PROF, quality ratings are not required at the item level.

In examining the mean values for each therapy domain, the results indicated that the Child therapy domain showed a larger effect size in its statistical significance between intervention videos in comparison to Context therapy. This finding was surprising. The Context domain scores were expected to be higher than the Child therapy domain within Context focused therapy videos because therapists were using a Context therapy approach on video. Using reverse scoring, Context therapy scores (2.69) were statistically significantly lower than Child therapy means (3.57), an unexpected finding because the scores were expected to be similar. This supports the hypothesis, expressed above; that it may have been more difficult for the interventionists on video to deliver and/or demonstrate the novel Context focused therapy. Alternatively, lower frequencies of Context therapy actions or behaviours could be appropriate for the type of intervention. One of the assumptions of this fidelity measure was that higher frequency

scores equals higher fidelity, but it did not take into consideration how to account for "appropriateness" of the frequency (or infrequency) of the behaviour. The context therapy data highlighted this challenge whereby the frequency scores were low yet the quality ratings were high. On the Child focused therapy videos, therapy domains showed more distinct means (i.e., larger gap between means), supporting the assumption that higher frequencies of child therapy measure higher fidelity. It is also plausible that child intervention strategies are more discrete or are easier to count. Similar results were evident in the clinical chart note data (frequency counts), whereby correlations with video frequency scores were moderate for Child therapy (r = .57) and very low for Context therapy (r = .13).

Global quality scores had high Inter-rater reliability and were highly correlated with quality domain scores, supporting the use of one quality rating for each video. Moving forward, examining quality scores for each domain could be done to determine if distinction between ratings for each therapy domain are feasible. Results for global fidelity scores demonstrated how blinded raters were still able to appropriately identify the type of intervention delivered on video. In future use of the PROF, this can be examined further as it may be more important that the raters are knowledgeable about the delivered interventions.

Relevance of Implementation Fidelity Framework

The PROF has successfully operationalized the important concepts of adherence and moderating factors from the IFF model. Characteristics of the model that the PROF has addressed include adherence of the intervention "content", including differentiation

between interventions through reverse scoring of the context and child therapy domains. The IFF adherence concept of "frequency" is evident in the PROF as the basis of a scoring scale. One issue for future clarification is the appropriateness of behaviour frequencies for some interventions, to answer the question whether more frequent observations of the behaviour is an indication of higher fidelity. This issue can be addressed within the moderating concept of "comprehensiveness of policy description", which refers to detailing the complexity and clarity of the intervention(s) (Dusenbury, Brannigan, Falco, & Hansen, 2003). This issue was not addressed in this study. Moderating factors of the IFF that are measured by the PROF include the novel concepts of "quality of delivery" and "client responsiveness". Rating the quality of therapist behaviours was evident in the item scaling and global quality rating. The concept of client responsiveness was present in the measure by evaluating child and parent behaviours, in addition to their interactions with the therapist.

Study Limitations

Several limitations were evident in this study. Therapy domains were validated by one external source (i.e., Chart Notes) with a small sample size (n=23). This may have impacted the correlations with the video ratings; however, the ratings for each domain within each video type demonstrated consistent patterns with video scoring. In addition, the general attributes domain was unable to be validated with an external source because of the lack of such tools in paediatric rehabilitation. Even though this general attributes domain was content validated through the reported Delphi study, future work could

examine behavior change, or therapeutic alliance measures in the psychology literature that may possess similar item constructs as described in Di Rezze et al., (2012).

Another limitation of this study is the cross-sectional nature of the data. Fidelity scores are meant to identify whether the characteristics of a therapy session demonstrate fidelity throughout the study. Examining multiple performances from each therapist across several clients could provide data to calculate a typical therapy session fidelity score. In this way, one could avoid the assumption that behaviours on one occasion would be similar across time and other children.

Based on the endorsement scores that show the percentage of items that were not observed, two items in the general attributes domain were absent from over 50% of the responses for both intervention types. These items were related to goal setting and whether the child experienced discomfort due to a personal issue. Since these videos were secondary data, they may have only provided the essential characteristics of the intervention delivered and not information on all general attributes. In future use of this measure, investigators could implement video procedures to ensure that the therapists video record all relevant aspects of the intervention session.

Implications

The PROF is a fidelity measure that takes approximately 10 hours for rater training within the context of an intervention study. Training and practice is likely necessary for any fidelity measure, and the time necessary for raters to be prepared to use the PROF is less than other fidelity measures in the literature, e.g., 44 hours (Martino et al., 2008). Based on the D-studies of each of the domains, good reliability scores

(minimum of .71) for the PROF can be achieved in future work involving five raters with similar background and training. Extra training for PT raters on the Context therapy approach will likely be necessary to ensure their ability to consistently rate context items. While this still represents a substantial time investment, fidelity data is vitally important in clinical trials.

Rating times for the measure will vary based on the length of the video, but without rating quality for each item, the time to rate each video would be significantly less than the average time of 17 minutes. With more clearly defined video procedures, the reliability scores could improve and a reduced number of raters would be required. More specific directions to ensure that the intervention session is continuously recorded will enable the observer to follow the entirety of the intervention. Improving the uniformity of the therapy videos will likely improve the psychometric properties of the PROF.

Changes to the PROF from the current findings suggest this measure will require more psychometric testing to refine its utility. Specific testing can involve examining: rater scores between balanced groups of OTs and PTs; quality ratings for each therapy domain; ratings of videos with more direct video procedures for the therapist delivering intervention; and the feasibility of multiple session ratings. Within the twenty items in the General Attributes domain, multiple persons were rated, such as the therapist, child, parent, and their interactions. Observed differences in the participation of each person were evident in the interventions (e.g., parent behaviours higher in context therapy and therapist behaviour potentially higher for child therapy). This finding raises additional questions about how the General Attributes domain is scored and whether sub-domains

might provide more insight into intervention behaviours. For example, sub-domains could be developed based on the person being observed (i.e., therapist, child, or parent). Another way to potentially create sub-domains could be based on literature that categorizes general attributes related to behaviour change (e.g., praise to re-enforce a behaviour) (Abraham & Michie, 2008), or Family-centred practice constructs related to interpersonal (e.g., warmth) and procedural characteristics (e.g., goal setting) (Dunst, 2009). A larger sample size is required to explore the relationship of each of these potential sub-domains using an Exploratory Factor Analysis.

For intervention research, the PROF can be utilized for this population with the same interventions, while the general attributes domain can be used for studies involving this population for other therapies. Furthermore, investigators of other interventions could use the framework of the PROF (i.e., frequency and quality scales) and its procedure (e.g., training process, rating process) as a template to develop unique therapy domains. The PROF provides paediatric rehabilitation investigators with a foundation on which to base their fidelity procedures within their program of research.

The work reported here is the beginning of research that will continue to test the measure and define its application in outcomes research. To identify the properties of the PROF as it relates to outcomes, future work will need to examine the relationship between observed items and their appropriate frequencies for specific outcomes. In the literature, studies describing novel fidelity measures have not traditionally reported cut-scores that identify an acceptable fidelity rating for the interventions delivered or the general attributes. The examination of fidelity cut-scores can be explored in a subsequent

study. Benchmarks of fidelity for each of the PROF domains will need to be identified to determine the standard of fidelity scores across the rated videos (i.e., what is acceptable and unacceptable fidelity). The outcomes in relation to fidelity scores will then address how the general attributes domain and therapy domains relate to outcomes. This knowledge can help to determine what domains are strong mediators or moderators of paediatric outcomes**.** The PROF may also be useful in the training and/or evaluating clinicians on novel therapies within practice, as well as an education tool for student clinicians within OT and PT programs.

References

Abraham, C., & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychology*, 27, 379-387.

Barber, J.P., & Crits-Christoph, P. (1996). Development of a therapist adherence/competence rating scale for supportive-expressive dynamic psychotherapy: A preliminary report. *Psychotherapy Research*, 6, 81-94.

Baron, R.M., & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.

Bayona, C.L., McDougall, J., Tucker, M.A., Nichols, M., & Mandich, A. (2006). School-based occupational therapy for children with fine motor difficulties: Evaluating functional outcomes and fidelity of services. *Physical & Occupational Therapy in Pediatrics, 26,* 89-110.

Bond, G.R., Campbell, K., Evans, L.J., Gervey, R., Pascaris, A., Tice, S., et al. (2002). A scale to measure quality of supported employment for persons with severe mental illness. *J Vocation Rehab, 17,* 239-250.

Boutron, I., Moher, D., Altman, D.G., Schultz, K.F., & Ravaud, P. (2008). Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment: Explanation and Elaboration. *Annals of Internal Medicine, 148,* 295-306.

Breitenstein, S.M., Gross, D., Garvey, C.A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation Fidelity in Community-Based Interventions. *Research in Nursing & Health, 33,* 164-173.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, A., Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, 2:40, doi: 10.1186/1748-5908-2-40

Carroll, K.M., Connors, G.J., Cooney, N.L., DiClemente, C.C., Donovan, D.M., Kadden, R.R., et al. (1998). Internal validity of project MATCH treatments: discriminability and integrity. *Journal of Consulting and Clinical Psychology, 66,* 290-303.

Carroll, K.M., Nich, C., Sifry, R.L., Nuro, K.F., Frankforter, T.L., Ball, S.A. (2000). A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug & Alcohol Dependence, 57,* 225-38.

Carroll, K.M., & Nuro, K.F. (2002). One size cannot fit all: A stage model for psychotherapy manual development. *Clinical Psychology: Science and Practice, 9,* 396-406.

Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284-290.

Darrah, J., Law, M., & Pollock, N. (2001). Family-centred functional therapy – A choice for children with motor dysfunction. *Infants and Young Children, 13,* 4, 79-87.

Darrah, J., Law, M.C., Pollock, N., Wilson, B., Russell, D.J., Walter, S.D., Rosenbaum, P., Galuppi, B. (2011). Context therapy: a new intervention approach for children with cerebral palsy. *Developmental Medicine & Child Neurology, 53,* 615–620.

Diamond, G.M., Diamond, G.S., & Hogue, A. (2007). Attachment-based family therapy: Adherence and differentiation. *Journal of Marital and Family Therapy, 33,* 2, 177-191.

Di Rezze, B., Law, M., Gorter, J.W., Eva, K., and Pollock, N., (In Press). A narrative review of generic intervention fidelity measures. *Physical and Occupational Therapy in Pediatrics.*

Di Rezze, B., Law, M., Eva, K., Pollock, N., and Gorter, J.W. *General Therapy Behaviours in Paediatric Rehabilitation: Essential Attributes for Intervention with Children with Physical Disabilities.* Manuscript in preparation.

Dunst, C.J. (2009). Implications of Evidence-Based Practices for Personnel Preparation Development in Early Childhood Intervention. *Infants and Young Children, 22,* 44-53.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. (2003) A Review of Research on Fidelity of Implementation: Implications for Drug Abuse Prevention in School Settings. *Health Education Research 18,* 237-256.

Eames, C., Daley, D., Hutchings, J., Hughes, J.C., Jones, K., Martin, P., Bywater, T. (2008). The leader observation tool: a process skills treatment fidelity measure for the Incredible Years parenting programme. *Child: care, health and development, 34,* 391-400.

Gardner, F. (2000). Methodological issues in the direct observation of parent-child interactions: do observational findings reflect the natural behaviour of participants? *Clin Child Fam Psychol Rev, 3,* 185-198.

Godfrey, E., Chalder, T., Ridsdale, L., Seed, P., & Ogden, J. (2007). Investigating the 'active ingredients' of cognitive behaviour therapy and counseling for patients with chronic fatigue in primary care: Developing a new process measure to assess treatment fidelity and predict outcome. *British Journal of Clinical Psychology, 46,* 253-272.

Hennessey, M.L. & Rumrill, P.D. (2003). Treatment fidelity in rehabilitation research. *Journal of Vocational Rehabilitation, 19,* 123-126.

Hogue, A., Liddle, H.A., Singer, A., & Leckrone, J. (2005). Intervention fidelity in family-based prevention counseling for adolescent problem behaviors. *Journal of Community Psychology, 33,* 2, 191-211.

Hollon, S.D. (1984). Final Report: System for rating psychotherapy audiotapes. Rockville, MD: Department of Health and Human Services.

King, G., Currie, M., Bartlett, D., Strachan, D., Tucker, M.A., & Willoughby, C. (2008). The development of expertise in paediatric rehabilitation therapists: the roles of motivation, openness, to experience, and types of caseload experience. *Australian Occupational Therapy Journal, 55,* 108-122.

Kraemer, H., Wilson, T., Fairburn, C., & Agras, W.S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59,* 877-883.

Law, M., Darrah, J., Pollock, N., Rosenbaum, P., Russell, D., Walter, S.D., Petrenchik, T., Wilson, B. & Wright, V. (2007). Focus on Function – a randomized controlled

trial comparing two rehabilitation interventions for young children with cerebral palsy. *BMC Pediatrics*, **7:**31doi:10.1186/1471-2431-7-31.

Law, M., Darrah, J., Pollock, N., King, G., Rosenbaum, P., Russell, D., et al. (1998). Family-centred functional therapy for children with cerebral palsy: An emerging practice model. *Physical and Occupational Therapy in Pediatrics, 18,* 1, 83-102.

Law, M., Darrah, J., Pollock, N., Wilson, B., Russell, D., Walter, S., Rosenbaum, P., Galuppi, B. (2011). Focus on function: a cluster, randomized controlled trial comparing child- versus context-focused intervention for young children with cerebral palsy. *Developmental Medicine and Child Neurology,* 53, 621-629.

Lucca, A.M. (2000). A clubhouse use fidelity index: Preliminary reliability and validity results. *Mental Health Services Research, 2,* 89-94.

Martino, S., Ball, S.A., Nich, C., Frankforter, T.L., & Carroll, K.M. (2008). Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence, 96,* 37-48.

McIntosh, V.V.W., Jordan, J., McKenzie, J.M., Lutty, S.E., Carter, F.A., Carter, J.D., Frampton, C.M.A., & Joyce, P.R. (2005). Measuring therapist adherence in psychotherapy for anorexia nervosa: Scale adaptation, psychometric properties, and distinguishing psychotherapies. *Psychotherapy Research, 15,* 3, 339-344.

Moras, K., & Hill, C.E. (1991). Rater selection in psychotherapy process research: Observations on the state-of-the-art. *Psychotherapy Research, 1*, 114-124.

Mowbray, C.T., Holter, M.C., Teague, G.B., and Bybee, D. (2003). Fidelity criteria: development, measurement and validation. *American Journal of Evaluation, 24,* 315-340.

Norman, G.R., and Streiner, D.L. (2008). *Biostatistics: The Bare Essentials 3rd Edition.* Hamilton: B.C. Decker Inc.

Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E., & Galuppi, B. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine and Child Neurology, 39,* 214-223.

Parham, L.D., Cohn, E.S., Spitzer, S., Koomar, J.A., Miller, L.J., Burke, J.P. et al. (2007). Fidelity in sensory integration intervention research. *The American Journal of Occupational Therapy, 61,* 216-227.

Parham, D., Roley, S., May-Benson, T., Koomar, J., Brett-Green, B., Burke, J.P., Cohn, E., Mailloux, Z., Miller, L.J., Schaaf, R., (2011). Development of a fidelity measure for research on the effectiveness of the Ayres Sensory Integration intervention. *American Journal of Occupational Therapy, 65,* 133-142

Perepletchikova, F., Treat, T.A., & Kazdin, A.E. (2007). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12,* 315-340.

Schlosser, R.W. (2002). On the importance of being earnest about treatment integrity. *Augmentative and Alternative Communication, 18,* 36-44.

Stein, K.F., Sargent, J.T., & Nicholas, R. (2007). Intervention research: Establishing fidelity of the independent variable in nursing clinical trials. *Nursing Research, 56,* 54-62.

Stratford, P.W., & Spadoni, G.F. (2003). Sample size estimation for the comparison of competing measures' reliability coefficients. *Physiotherapy Canada, 55,* 225-229.

Streiner, D.L. & Norman, G.R. (2006). *Health Measurement Scales: A practical guide to their development and use 3rd Ed.* New York: Oxford University Press.

Waltz, J., Addis, M.E., Koerner, K. & Jacobson, N.S. (1993). Testing the integrity of psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and clinical psychology, 61,* 620-630.

Webb, C.A., DeRubeis, R.J., and Barber, J.P. (2010). Therapist adherence/competence and treatment outcome: a meta-analytic review. *Journal of Consulting and Clinical Psychology, 78,* 200-211.

Table 1. Mean Rater Scores for second viewing of video to test rater drift

Rater	Child Domain (Video 1 = 4.0)	Context Domain (Video 1 = 2.29)	Gen Attributes Domain (Video 1 = 3.37)
1	3.75	2.25	3.4
2	4	2.25	3.4
3	2.75	3.25	3.5
4	4.25	2.25	3.55
5	3.75	2.25	3.35
6	3	1.75	2.95

Table 2. Descriptive data for each domain within each round of video rating

The Round each rater viewed videos	Domain	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	Child	4.31	.32	3.68	4.95
	Context	2.19	.27	1.64	2.73
	General Attributes	3.46	.15	3.16	3.76
2	Child	3.46	.16	3.14	3.77
	Context	2.54	.14	2.27	2.80
	General Attributes	3.21	.07	3.07	3.36
3	Child	3.40	.15	3.11	3.69
	Context	2.51	.13	2.26	2.76
	General Attributes	3.22	.07	3.08	3.36
4	Child	3.47	.15	3.18	3.77
	Context	2.43	.13	2.17	2.68
	General Attributes	3.20	.07	3.06	3.33

Table 3. Descriptive data for each domain and rater profession

Rater	Domain	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
OT	Child	3.62	.10	3.42	3.82
	Context	2.40	.09	2.22	2.59
	General Attributes	3.19	.06	3.08	3.30
PT	Child	3.25	.14	2.97	3.54
	Context	2.58	.13	2.31	2.85
	General Attributes	3.30	.08	3.15	3.46

Table 4. Reliability data for frequency and quality scores across domains

	Frequency Scale		**Quality Scale**	
Domain	**Alpha**	**Inter-rater reliability G(6) [G(1)]**	**Alpha**	**Inter-rater reliability G(6) [G(1)]**
Child Therapy	.85 (k=4)	.87 [.54]	.86 (k=4)	.41 [.10]
Context Therapy	.71 (k=4)	.75 [.33]	.90 (k=4)	.41 [.10]
General Attributes	.78 (k=20)	.82 [.44]	.81 (k=8)	.41 [.10]

Note. k=number of items included within each analysis; G = G-coefficient; G-coefficients in [ ] indicate single rater scores

Inter-rater Reliability

Table 5. Descriptive data and Inter-rater reliability of all domains by intervention video type for frequency and quality ratings.

	Frequency Ratings				**Quality Ratings**			
	Child Video		Context Video		Child Video		Context Video	
Domain	Mean(SD)	IRR (6 raters)	Mean(SD)	IRR (6 raters)	Mean(SD)	IRR (6 raters)	Mean(SD)	IRR (6 raters)
Child Therapy	4.10 (.68)	**.40**	2.76 (.89)	**.78**	3.87 (.72)	**0.33**	3.40 (.77)	**0.62**
Context Therapy	2.08 (.86)	**.55**	2.95 (.83)	**.70**	3.45 (.80)	**0.79**	3.68 (.85)	**0.60**
General Attribute	3.26 (.51)	**.57**	3.18 (.61)	**.80**	3.75(.70)	**0.35**	3.50 (.66)	**0.30**

Table 6. Pearson's Correlation coefficients of relationships between frequency for each domain

Domain	Child Therapy	Context Therapy	General Attributes
Child Therapy	1.00	N	P
Context Therapy	-0.71**	1.00	P
General Attributes	0.43*	-0.07	1.00

Note. ** indicates p=0.01; * indicates p=0.05; N= Negative correlation; P=positive correlation

Table 7. Frequencies of child and context behaviours from clinical charts for each intervention type

	Clinical Chart Notes	
Video Type	**Child**	**Context**
Child Therapy	23	5
Context Therapy	0	14

Table 8. Pearson's correlation between coded clinical charts and frequency scores of videos across both therapy domains

		Video Frequencies	
	Domains	Child	Context
Chart Frequencies	Child	0.57**	-0.47*
	Context	-0.22	0.13

Note. ** indicates p=0.01; * indicates p=0.05

Table 9. One-way ANOVA for each rating type and domain by intervention type

Frequency Ratings				
Domain	**Child Therapy Video Mean (SD)**	**Context Therapy Video Mean (SD)**	**F-ratio**	**p-value**
Child Therapy domain	4.10(.68)	2.76(.89)	108.26	<.001
Context Therapy domain	2.08(.86)	2.95(.83)	38.96	<.001
General Attributes domain	3.26(.51)	3.18(.61)	0.65	<.420
Quality Ratings				
Domain	**Child Therapy Video Mean (SD)**	**Context Therapy Video Mean (SD)**	**F-ratio**	**p-value**
Child Therapy domain	3.79 (.76)	3.42 (.76)	8.777	.004
Context Therapy domain	3.45 (.80)	3.69 (.83)	3.124	.079
General Attributes domain	3.69 (.74)	3.50 (.66)	2.551	.112

Figure 1. The Implementation Fidelity Framework (Carroll et al., 2007)

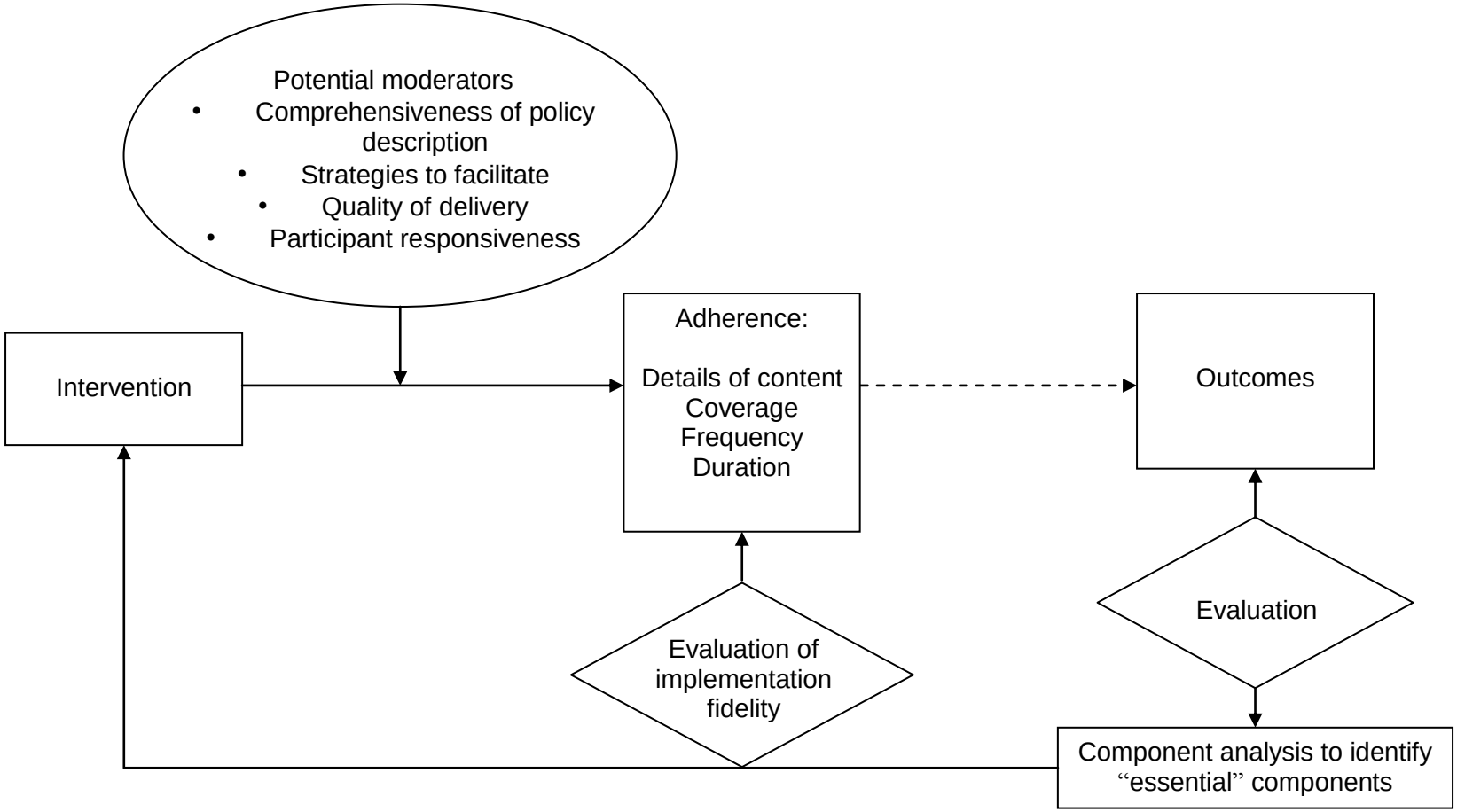
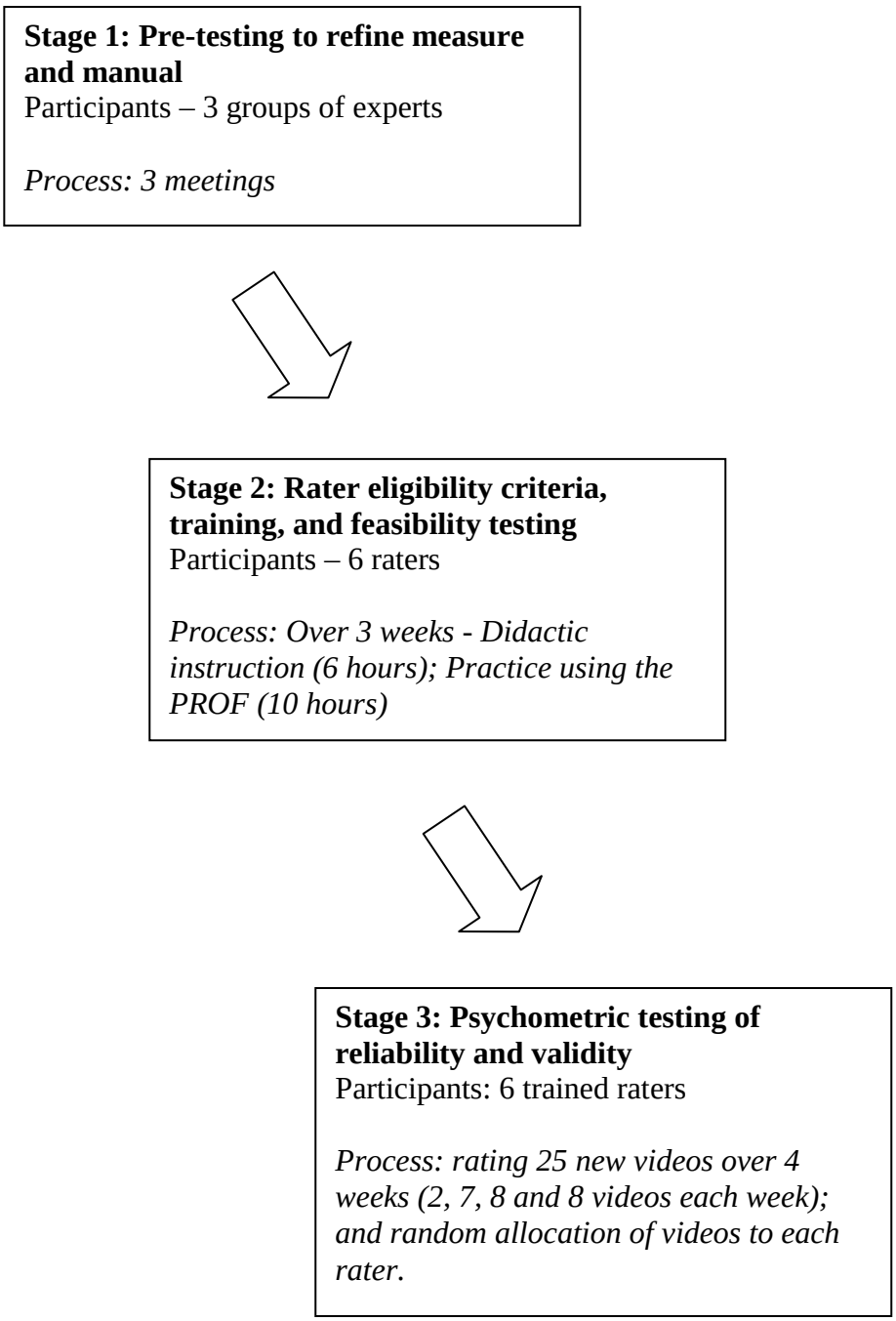


Figure 2. Scales used for the intervention fidelity scale

Frequency Rating	N/ A Not Present	1 Not at all	2 A little	3 Somewhat	4 Considerably	5 Extensively
Quality Rating	N/A Not Applicable	1 Poor	2 Acceptable	3 Average	4 Good	5 Very Good

Figure 3. Flow Chart of the stages to develop the PROF



Chapter Five: Discussion

Over the last 20 years, research examining the impact of rehabilitation interventions has increased and there is now more knowledge regarding the efficacy of specific rehabilitation interventions. While these developments are very positive, this evidence has not been sufficient to identify "how and why it works, when and how much is best, and for whom" (Hart, 2009, p. 825). In many rehabilitation studies, both for children and adults, specific information about the intervention and its delivery are not always provided and there are few studies examining the relationship between the implementation of intervention strategies and outcomes. This concept, known as intervention fidelity, examines to what extent an intervention was implemented as planned or delivered based on its foundational principles (Waltz, Addis, Koerner, & Jacobson, 1993). To better understand the mechanism of the intervention that influences outcomes, it is essential that the intervention be clearly defined, measurable and delivered appropriately. Knowledge about the impact of specific actions from an intervention on outcomes will guide other therapists in their treatment for similar clients.

In rehabilitation, the process of intervention is complex because therapists deliver an intervention guided by general (or non-specific) therapy approaches (e.g., goal setting, grading difficulty, rapport building) as well as utilizing specific treatment strategies for the therapy. In paediatric rehabilitation, intervention is further complicated by the need to consider the child's developmental level, likes/dislikes (for motivation), and other needs involving both the child and the family (Lane & Bundy, 2012). These factors have made

it challenging for investigators in paediatric rehabilitation to identify general therapy approaches that are observable and measurable.

Psychology researchers have had some success in measuring general therapy behaviours, labeled therapeutic alliance, and defined as the interpersonal characteristics between the client and therapist (Elvins & Green, 2008). Therapeutic alliance behaviours have demonstrated high predictive validity in relation to outcomes in adult literature (Green, 2006). On the other hand, the field of therapeutic alliance has been understudied in child and adolescent psychology research (Green, 2006). Similarly, there is very little research in paediatric rehabilitation that has studied non-specific intervention behaviours demonstrated by the child, therapist, and parent in the context of Family-Centred Services (FCS). Examining both the intervention-specific and non-specific behaviours has the potential to provide an improved understanding of the key mechanisms of therapy on outcomes. The development of a fidelity measure to evaluate these behaviours of the therapist, the child and the parent is an important starting point to begin this investigation in the field of paediatric rehabilitation.

Most instruments measuring intervention fidelity in efficacy research have been defined as intervention-specific measures. These measures examine the active ingredients of a particular intervention being studied; however, ‘generic’ fidelity measures examine the characteristics of more than one intervention and characteristics common to all interventions (e.g., general therapy behaviours) (Brientenstein et al., 2010). Generic measures of intervention fidelity identify the presence of all relevant therapy behaviours that may impact on outcomes. Such measures allow comparison across different

treatments. This type of fidelity measure is capable of measuring the elements of complex therapies in a trial. Traditional fidelity measures do not capture this range of therapy behaviours as they only measure strategies specific to one intervention. Currently, the only type of intervention session-focused fidelity measures in the paediatric rehabilitation literature are traditional, intervention-specific measures. Knowledge of the presence of intervention specific and non-specific behaviours can enable an investigator to examine whether there is an interdependence relationship between them, and to what extent the general therapy behaviours impact on client outcomes.

The objectives of this thesis were to examine the conceptual foundations of fidelity measurement and intervention that best align with paediatric rehabilitation, and to operationalize these key behaviours within a novel generic fidelity measure. The results from this thesis provide evidence for the use of generic fidelity measures and their value in identifying non-specific therapy behaviours that include observations of the child and parent. In addition, this measure emphasizes the importance of examining the behaviours of multiple interventions to identify if, or to what degree, contamination has occurred in the way in which therapy is delivered in a study.

Through a series of three studies, this thesis has resulted in the development of the first generic fidelity measure for paediatric rehabilitation - the Paediatric Rehabilitation Observational measure of Fidelity (PROF). The PROF was developed through foundational studies that have made novel contributions to the field of fidelity measurement and paediatric rehabilitation. The narrative literature review (Chapter 2) provided a detailed analysis of generic fidelity measures from psychotherapy and their

relevance to paediatric rehabilitation. In Chapter 3, the Delphi study described the results of a consensus exercise with experts to identify the essential general therapy attributes of paediatric rehabilitation intervention for children with physical disabilities. These attributes formed the basis of the items in the general therapy domain in the PROF. The final study (Chapter 4) reported the development and psychometric testing of the PROF itself. A summary of the findings for each chapter is below.

Chapter 2

The purpose of Chapter 2 was to describe the characteristics of existing generic fidelity measures and examine how these attributes fit with paediatric rehabilitation. The narrative review reported, for the first time, an in-depth analysis of the characteristics of generic fidelity measures. The term, "generic fidelity measure" has only recently been described by Breitenstein and colleagues (2010). In the literature, there are many descriptive papers that identify conceptual models and steps to create fidelity measures that are only relevant to a single, specific intervention approach. Thus, the narrative review guided the development of the PROF and provided detailed information about the common procedures and characteristics of generic fidelity measures to examine multiple interventions in a randomized controlled trial (RCT). Examples of these procedures included what intervention behaviours to measure (e.g., measurement of frequency and quality of behaviours), as well as the process to train raters to accurately and consistently use the measure with videotaped intervention sessions. Furthermore, the psychometric properties of these measures were summarized to provide standards achieved by other complex interventions. These data provide researchers with an understanding of the

emerging factors to consider in the process of fidelity measurement for complex, non-pharmacological interventions.

In Chapter 2, the key findings focused on the identification of item content for the general therapy behaviours non-specific to an intervention. The review examined the relevance of 76 general therapy items to the Family-Centred Service (FCS) framework that is used extensively in paediatric rehabilitation (Rosenbaum et al., 1998). Of these items, thirty seven aligned with the FCS and included observable behaviours of therapist, client and therapist-client interactions. These results have initiated the dialogue regarding important, general therapy behaviours within an intervention session that are observed across paediatric rehabilitation interventions that use a family-centred services framework.

Chapter 3

In Chapter 3, a Delphi study was conducted to generate attributes of general therapy behaviours essential to family-centred service for children with physical disabilities. The Delphi study involved an exercise with an interdisciplinary group of health research experts to identify a set of general therapy attributes in paediatric rehabilitation for children with physical disabilities. Thirty-five attributes were generated over the course of three rounds of the Delphi process. These thirty-five general therapy behaviours were characterized by categories related to actions of the therapist (21), client (9) and interactions between the client and therapist (5). Client behaviours included involvement of either the child or the parent within the intervention. Examples of attributes relevant for each category were: Therapist compliments and/or praises the

child's effort or performance within the intervention session; child/parent communicates his/her preference(s) in the intervention session; and therapist and child/parent show a strong rapport in the intervention session. Concurrently, an independent data source (i.e., content analysis of therapist interviews) aimed to be a supplementary source of general attributes. The results of this qualitative content analysis indicated that 17/19 of which 15 are useable attributes corroborated with the general therapy behaviours from the Delphi Process.

This work has, for the first time, defined the general therapy behaviours of paediatric rehabilitation deemed essential within an intervention session. These behaviours have a potentially important influence on outcomes and can serve as a basis of observable and measurable general therapy behaviours. The relevance of these attributes across other paediatric disability populations is yet to be examined, and this work serves as the foundation for future work for others who are delivering family-centred services.

Chapter 4

Following the identification of general attributes, a novel fidelity measure to assess the specific and general therapy behaviours was developed. The study of the development and psychometric properties of the PROF reported the first paediatric rehabilitation observational fidelity measure that evaluates multiple unique interventions and items common to both interventions. The measure was tested using data from a completed randomized controlled trial with young children with cerebral palsy (Law et al., 2011). The PROF demonstrated sound reliability properties (G-coefficients > 0.7) for both intervention specific and non-specific domains. The level of reliability for the PROF

in this study was comparable to the properties of existing generic fidelity measures (Hogue, Liddle, Singer, & Leckrone, 2005; Martino, Ball, Nich, Frankforter, & Carroll, 2008). In terms of validity, the study demonstrated how each scaling construct of the PROF (frequency vs. quality) and each domain (child therapy vs. context therapy vs. general therapy behaviours) were distinct from one another.

This research demonstrates the ability to generate a reliable and valid generic fidelity measure for intervention involving children with cerebral palsy. More specifically, it highlights the ability to observe and measure non-specific intervention items. The PROF has been grounded in a recent conceptual model, the Implementation Fidelity Framework (Carroll et al., 2007), that examines intervention behaviours of the therapist (i.e., quality of delivery) and responsiveness of the client. Furthermore, the scaling characteristics of the PROF differ from any previous fidelity measure in paediatric rehabilitation, whereby therapist behaviours were rated in terms of both frequency and quality of intervention delivery.

Contributions of this Research to Paediatric Rehabilitation

This thesis has resulted in several key contributions to intervention research within the field of paediatric rehabilitation. First, this work has more broadly defined the potential active ingredients of paediatric rehabilitation to include general therapy behaviours and demonstrated that these behaviours were measurable. Second, this thesis included the development of a fidelity measure based on a current conceptual model of fidelity. Third, a novel analysis was conducted to combine the therapists' performance for each intervention to calculate an overall fidelity score. Finally, this thesis has contributed

significantly to the area of cerebral palsy research by developing a fidelity measure that will be useful to test two important interventions used by occupational therapists (OT) and physiotherapists (PT).

Identifying measurable general therapy behaviours

The active ingredients of intervention in fidelity measures have traditionally focused solely on the specific content of the therapy delivered, and demonstrated solely by the therapist. The characteristics of general therapy behaviours have been judged to be important but difficult to measure because of the challenges in standardizing paediatric intervention and a readily quantifiable process in discrete, consistent units (Michaud, 2004). The individualized process of assessment and intervention in paediatric rehabilitation are essential to family-centred services where therapists and the family work together as partners to define the therapeutic needs of the child (Law, Missiuna, Pollock and Stewart, 2001). Furthermore, the therapeutic process requires tailoring of the intervention to the particular skills, needs, and interests of the client and the "degree to which and the finesse with which therapists" are able to be successful in this process varies by therapist and from situation to situation (Mattingly & Fleming, 1994, p.338). Key factors such as child development and varying ability contribute to this challenge of individualizing intervention.

This thesis has identified both attributes and measurable items to describe the general therapy behaviours identified as important characteristics of paediatric rehabilitation, in particular FCS. To date, FCS research has examined the principles and attributes of health care delivered [at a service level, using the Measure of Processes of

Care (King, Rosenbaum & King, 1995)], but research has yet to measure the therapy process behaviours within an intervention session. This thesis has started to fill this research gap by generating observable general therapy behaviour items within the PROF.

The items in the PROF have demonstrated strong psychometric properties across a wide range of demographic characteristics of children within the intervention videos. Variation in the developmental level and abilities of the child included a range of ages 1 to 6 years (mean = 3.9, SD = 1.3) and child gross motor abilities Gross Motor Function Classification System (Palisano et al., 1997), levels I to III, 60% (ambulatory); and levels IV and V, 40% (non-ambulatory). Parent presence on the video also varied with 36% of parents in attendance within the therapy session. The therapists delivering the interventions within the videos varied in terms of profession, with 56% being OT and 44% PT. The fact that the PROF possessed good reliability and validity ratings across such variability indicated that these general therapy items were observable and can be objectively rated.

Utilizing a conceptual foundation for fidelity measurement

Conceptual models assist with the operationalization of theory to define key concepts and their interrelationships (Whyte, 2008). Thus, using conceptual models for developing a fidelity measure can explain the characteristics that the measure could examine. In the literature, fidelity measures often lack a foundational model of fidelity that relates to the appropriate intervention(s) involved in the study. This thesis has developed a generic fidelity measure grounded within a recent conceptual framework aligned with paediatric rehabilitation – the Implementation Fidelity Framework (IFF).

The IFF emphasizes the importance of clear **definition** of the study interventions, the evaluation of potential moderators of intervention adherence, such as **general therapy attributes** (i.e., client responsiveness) and the measurement of the **quality of therapy delivery**. The advantage of using the IFF is that it incorporates all aspects of complex intervention for the fidelity measure, including those that may be overlooked in typical intervention-specific measures.

Through the evaluation of client responsiveness within the PROF, the IFF guided the integration of the concept of client adherence into the examination of intervention fidelity. This nuance of fidelity measurement recognizes that complex therapies have a dynamic nature, whereby the client's performance can influence the way in which therapy is delivered. Thus, the inclusion of the evaluation of child and parent behaviours within the PROF has taken into consideration potential key moderators of intervention delivery in paediatric rehabilitation.

In the PROF, the use of a scale that evaluates quality of therapist performance was unique to existing fidelity measures in paediatric rehabilitation. This scale provides further insight regarding the frequency scores to ensure that that the performance of behaviours were 'meaningfully' demonstrated, rather than being delivered in an inappropriate manner (Carroll et al., 2000). The findings indicated some difficulties in the PROF in the consistent rating of quality across all items (G =.41), but found very good reliability of quality ratings at a global item level (G=.80). These results represent a starting point to further explore how quality can be measured and levels of quality

defined within specific interventions and across general therapist behaviours within family-centred service.

Generating an overall fidelity score

As a generic fidelity measure, one of the characteristics of the PROF is that it possesses items differentiating the two interventions involved in the trial. The concept of therapy differentiation is defined as identifying the unique features of the interventions so that they can be reliably differentiated from each other (Dusenbury, Brannigan, Falco, & Hansen, 2003). For the observed intervention sessions, this differentiation sets the expectation that if the therapy was delivered with a high level of fidelity, then the scores will differ for the ratings of each unique intervention. Although the concepts of prescribed and proscribed behaviours have been discussed in the literature (Dane & Schneider, 1998; Dusenbury et al., 2003; Smith, Daunic, & Taylor, 2007), none of the identified measures from the narrative review (Di Rezze, Law, Gorter, Eva, & Pollock, In Press) have developed an analytic strategy to examine therapist behaviours of all interventions together. Thus, one novel contribution of the PROF to the field of fidelity measurement is the use of an analysis based on the concept of differentiation that combines intervention domain ratings with reverse scoring. Typically, generic fidelity tools have separately examined the fidelity scores for different interventions to identify the prescribed and proscribed behaviours. The reverse scoring analysis uses both intervention domains within the PROF to calculate an overall fidelity score examining all intervention content delivered by the therapist. In future work, this score can serve as the basis to identify a standard for an overall fidelity score. Such scoring will help identify

situations in which the fidelity of one intervention is negatively influenced by contamination from the other intervention.

Advancing the field of cerebral palsy intervention research

The PROF has also made a significant contribution in the area of cerebral palsy (CP) research by defining observable and measurable behaviours for two key interventions provided to children with CP by occupational and physiotherapists. This measure can now be used to examine the impact of each intervention, child or context focused therapy, on outcomes. This work has also identified an important question about the interventions - what is the appropriate level of frequency in the performed behaviour? In defining the characteristics of a novel intervention, knowledge about the level of acceptable frequency is important to understand. In Chapter 4, context focused therapy frequency ratings were lower than those of child focused therapy, but global quality ratings were similarly high. These results indicate that the low frequency of context focused therapy behaviours does not necessarily indicate that therapy was poorly and inappropriately delivered. As the intervention is novel, little is known about the expected frequency of the context therapy items in an ideal therapy session. This highlights the importance for investigators to consider both the active ingredients of an intervention as well as the meaningful frequency ratings for any emerging interventions. For context therapy, future examination of its delivery in other studies can provide additional knowledge regarding appropriate item frequency ratings and these levels can be described in the PROF manual.

Implications for using the PROF

The PROF has been developed to be used in future research trials involving child and/or context therapy interventions for children with CP. These interventions represent a predominant intervention (i.e., child focused therapy) and emerging intervention (i.e., context focused therapy) in paediatric rehabilitation for children with CP. Findings from this thesis support the use of the PROF in this manner. In particular, the PROF can be utilized to evaluate the specific item fidelity within replication studies using these interventions, such as the Learn 2 move 2-3 study (Ketelaar et al., 2010). The generic domain of the PROF has the potential to be used in other paediatric rehabilitation studies to examine general therapist and client behaviours. Before the PROF can be used in future studies, several recommendations from this research are proposed to describe how it can be used and how to enhance its success.

Using the PROF as a Framework

The PROF is a groundbreaking measure in paediatric rehabilitation that can serve as a framework for other investigators to develop future generic fidelity tools in this clinical area. Generic fidelity measures such as the YACS (Carroll et al., 2000) and the CSPRS (Hollon, 1984) are examples of tools in psychotherapy that have served as a basis for other versions that include other interventions (Martino et al., 2008) and/or population groups such as the CSPRS –HIV (Markowitz, Spielman, Scarvalone, & Perry, 2000).

The PROF measure provides researchers with procedures, training materials, a measurement scaling framework, and statistical methods to analyze its properties. The structure of the tool will enable future investigators to format novel intervention items

into new intervention domains. Investigators may also test existing general therapy behaviour items and/or trial other general attributes from Chapter 3 to determine the essential non-specific intervention behaviours. The procedures provided by the PROF include guidance on item development for specific interventions items and general therapy attribute items. This information can serve as a starting point for items related to other interventions. Materials from this research that provide methods to train raters, evaluate their performance, and approaches to remediation are also valuable resources for investigators. Finally, the PROF provides other family-centred interventions with examples of measurement scales related to frequency and quality and how to examine their psychometric properties. For any measure, the psychometric properties need to be tested in any new context or population. However, the measurement of the fidelity of general therapy behaviours in future research could benefit from the novel considerations raised in the development of the PROF as a detailed measure of family-centred rehabilitation behaviours for intervention sessions.

Using the PROF in a replication study

To use the PROF in future research, findings from this thesis show that some modifications to the rating scales, video procedures and raters are necessary to utilize it in its best state. The measure consists of 30 items, including two global items that have demonstrated good reliability for the frequency scale.

For ratings related to quality, findings indicate that global quality scores are the only reliable method to judge the overall quality therapy delivery in each video. If the quality scale is to be re-examined on an item level, it is recommended that clear examples

of high and low quality performance for each intervention be added to the PROF manual. These details may help to improve the interrater reliability of raters of quality scores at an item level, but the psychometric properties of these data would require re-examination.

Prior to intervention delivery within a future study, a clear procedure for videotaping each session will need to be provided to intervention therapists. The procedure should include a description of the expected content within the therapy session, suggestions regarding video camera set-up, and re-iteration of the therapy goals, type and timing of the intervention session. These recommendations will ensure that video content will be consistent across intervening therapists and can contribute to improved PROF psychometric properties.

Trained raters need not be blinded to the purpose of the fidelity exercise, but do need to be independent of the intervention study. For established interventions, these raters can be either OT or PT since both have demonstrated similar scoring behaviours in this situation. For novel interventions, such as the context therapy domain, it is important to ensure that both professions have a similar understanding of the observed behaviours. Discipline specific questionnaires regarding the therapists' knowledge about the intervention approaches could aid in this process.

Limitations

There are limitations within this research. The cross-sectional nature of the video data used to test the PROF restricts the judgment of the fidelity of therapist behaviours for a particular child and family to one point in time. A single time-point to observe and rate the fidelity of an intervention session with a specific child/family cannot be

extrapolated to another therapy session or family due to the influence of context specificity. Context specificity has been defined as the "observation that an individual's performance...in a particular situation is only weakly predictive of the same individual's performance...in a different situation" (Eva, 2003, p.587). Future work can use the PROF to examine the 'overall' fidelity of intervention sessions of therapists with different clients. This could provide information about the level of fidelity score that indicates a stable intervention over time and provide additional information regarding which variables contribute the most error to measurements collected.

A second limitation in this thesis is that the development of the PROF did not take into consideration all of the conceptual characteristics of the Implementation Fidelity Model (IFF) (Carroll et al., 2007). In the development of the PROF, two of the four potential moderator concepts were not addressed, the concepts of "comprehensiveness/complexity of policy description" and "strategies to facilitate implementation." Both of these concepts appear to represent content that was outside the scope of this thesis – to develop a measure to evaluate fidelity of an intervention session. The policy (or intervention) descriptions were already developed for the randomized trial and the strategies provided to the intervention therapists through training within the study. In future studies, integrating these concepts into the whole fidelity process can potentially help with challenges such as consistent video procedures, expected standards for frequency scores for the novel (i.e., context) therapy, and examples (strong and poor) of the expected quality of performance for each intervention.

Future Directions

The findings from this thesis mark the beginning of a series of inquiries that will provide insight and knowledge about fidelity measurement in paediatric rehabilitation. Future studies can examine judgments about what constitutes high/moderate/low fidelity scores in each domain for the interventions. This work can establish a standard of acceptable fidelity behaviours leading the way for investigators to study the relationship of fidelity scores to client outcomes of both child and context therapy.

Subsequent research could use knowledge about the intervention fidelity scores to identify behaviours that impact on intervention outcomes, and potentially determine the key ingredients that will have the most significant impact in efficacy studies. The association between ratings from the PROF and outcomes data will also identify whether the constructs are mediating or moderating factors through the use of more sophisticated statistical methods, such as hierarchical linear modeling.

One final research direction to be explored in the future is the meaningfulness of the general therapy attributes (from the Delphi study) within other child disability populations, other health disciplines, and other paediatric intervention approaches. General therapy attributes may also be examined in terms of relevance to core competencies of childhood interventions for clinicians and in-training students preparing for the field of family-centred service in childhood disability. Currently, there is a lack of knowledge regarding the attributes of general therapy across disciplines and types of clinical practice (Dunst et al., 2009).

Conclusions

This research has led to the development of the first generic fidelity measure in paediatric rehabilitation. To date, the use of a generic fidelity approach to examining intervention delivery has not been explored in paediatric rehabilitation. In the same way that appropriate outcome measures need to be considered in intervention studies, intervention research needs to consider appropriate fidelity measures to examine all salient characteristics of the intervention under investigation. Future research will result in further refinement of the PROF and its potential for use as a framework for other family-centred paediatric interventions across childhood disability populations. This thesis adds important knowledge to the limited evidence about how to accurately measure complex behaviours within family-centred paediatric rehabilitation. These complex behaviours include the range of essential behaviours within a therapy session by the therapist, child and parent, or their interactions. The examination of intervention fidelity is important for all research involving non-pharmacological interventions. Although the development and testing of any fidelity tool may require time and resources, a generic fidelity tool, such as the PROF, can reduce the amount of work for future research in similar areas of paediatric rehabilitation.

References

Breitenstein, S.M., Gross, D., Garvey, C.A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation Fidelity in Community-Based Interventions. *Research in Nursing & Health, 33,* 164-173.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science 2*:40 doi:10.1186/1748-5908-2-40

Carroll, K.M., Nich, C., Sifry, R.L., Nuro, K.F., Frankforter, T.L., Ball, S.A. (2000). A general system for evaluating therapist adherence and competence in psycho-therapy research in the addictions. Drug and Alcohol Dependence, 57, 225-238.

Coster, W., Tickle-Degnen, L., and Armenta, L. (1995). Therapist-child interaction during sensory integration treatment: Development and testing of a research tool. *The Occupational Therapy Journal of Research*, 15, 1, 17-35.

Dane, A., & Schneider, B. H. (1998). Integrity in primary prevention programs: Are implementation effects out of control? *Clinical Psychology Review, 18*, 23-45.

Di Rezze, B., Law, M., Gorter, J.W., Eva, K., and Pollock, N., (In Press). A narrative review of generic intervention fidelity measures. *Physical and Occupational Therapy in Pediatrics.*

Dunst, C.J. (2009). Implications of Evidence-Based Practices for Personnel Preparation Development in Early Childhood Intervention. *Infants and Young Children, 22,* 44-53.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. (2003) A Review of Research on Fidelity of Implementation : Implications for Drug Abuse Prevention in School Settings. *Health Education Research 18,* 237-256.

Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28,* 1167-1187.

Eva, K. (2003). On the generality of specificity. *Medical Education, 37,* 587-588.

Green, J. (2006). Annotation: The therapeutic alliance – a significant but neglected variable in child mental health treatment studies. *Journal of Child Psychology and Psychiatry, 47,* 425 - 435.

Hart, T. (2009). Treatment definition in complex rehabilitation interventions. *Neuropsychological Rehabilitation, 19,* 824-840.

Hogue, A., Liddle, H.A., Singer, A., & Leckrone, J. (2005). Intervention fidelity in family-based prevention counseling for adolescent problem behaviors. *Journal of Community Psychology, 33,* 2, 191-211.

Hollon, S.D. (1984). Final Report: System for rating psychotherapy audiotapes. Rockville, MD: Department of Health and Human Services.

Ketelaar, M., Kruijsen, A., Verschuren, O., Jongmans, M., Gorter, JW, Verheijden, J., Reinders-Messelink, H., Lindeman, E. (2010). LEARN 2 MOVE 2-3: a randomized controlled trial on the efficacy of child-focused intervention and context-focused intervention in preschool children with cerebral palsy. *BMC Pediatrics*, *10*:80. doi:10.1186/1471-2431-10-80

King, S., Rosenbaum, P., & King, G. (1995). The Measure of Processes of Care: A means to assess family-centred behaviours of health care providers. Hamilton, ON: McMaster University, Neurodevelopmental Clinical Research Unit.

Lane, S., and Bundy, A. (2012). Kids Can Be Kids: A Childhood Occupations Approach. Philadelphia, PA: F.A. Davis Company.

Law, M., Darrah, J., Pollock, N., Wilson, B., Russell, D., Walter, S., Rosenbaum, P., Galuppi, B. (2011). Focus on function: a cluster, randomized controlled trial comparing child- versus context-focused intervention for young children with cerebral palsy. *Developmental Medicine and Child Neurology,* 53, 621-629.

Law, M., Missiuna, C., Pollock, N., and Stewart, D. (2001). Foundations for Occupational Therapy Practice with Children. In J. Case-Smith (Ed.), *Occupational Therapy for Children.* St. Louis, MI: Mosby, Inc.

Markowitz, J.C., Spielman, L.A., Scarvalone, P.A., & Perry, S.W. (2000). Psychotherapy adherence of therapists treatment HIV-Positive patients with depressive symptoms. *Journal of Psychotherapy Practice Research, 9,* 2, 75-80.

Martino, S., Ball, S.A., Nich, C., Frankforter, T.L., & Carroll, K.M. (2008). Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence, 96,* 37-48.

Mattingly, C., and Fleming, M. (1994). *Clinical Reasoning: Forms of Inquiry in a Therapeutic Practice.* Philadelphia, PA: F.A. Davis Company.

Michaud, L. (2004). Prescribing therapy services for children with motor disabilities. *American Academy of Pediatrics, 113,* 1836-1838.

Norman, G.R., and Streiner, D.L. (2008). *Biostatistics: The Bare Essentials 3rd Edition.* Hamilton: B.C. Decker Inc.

Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E. & Galuppi, B. (1997). Development and validation of a gross motor function classification system for children with cerebral palsy. *Developmental Medicine and Child Neurology*, *39*, 214–223.

Rosenbaum, P., King, S., Law, M., King, G., and Evans, J. (1998). Family-Centred Service: A conceptual framework and research review. *Physical & Occupational Therapy in Pediatrics, 18,* 1, 1-20.

Smith, S.W., Daunic, A.P., Taylor, G.G. (2007). Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. *Education and Treatment of Children, 30,* 4, 121-134.

Waltz, J., Addis, M.E., Koerner, K. & Jacobson, N.S. (1993). Testing the integrity of psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61,* 620-630.

Whyte, J. (2008). A grand unified theory of rehabilitation (We Wish!). The 57th John Stanley Coulter Memorial Lecture. *Archives of Physical Medicine and Rehabilitation, 89,* 203-209.

Bibliography

Abraham, C., & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. Health Psychology, 27, 379-387.

Allen, K.D., Wilczynski, S. M., and Evans, J.H. (1997). Pediatric rehabilitation: Defining a field, a focus and a future. *International Journal of Rehabilitation and Health, 3,* 25-40.

Altman, D.G., Schulz, K,F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P.C., Lang, T.; CONSORT GROUP (Consolidated Standards of Reporting Trials). (2001). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine, 134,* 663-94.

Baer, J.S., Ball, S.A., Campbell, B.K., Miele, G.M., Schoener, E.P., and Tracy, K. (2007). Training and fidelity monitoring of behavioral interventions in multi-site addictions research. *Drug and Alcohol Dependence, 87*, 107-118.

Bartlett, D.J., Macnab, J., Macarthur, C., Mandich, A., Magill-Evans, J., Young, N.L., Beal, D., Conti-Becker, A., & Polatajko, H. (2006). Advancing rehabilitation research: an interactionist perspective to guide question and design. *Disability and rehabilitation*, *28*, 1169-76.

Bayona, C.L., McDougall, J., Tucker, M.A., Nichols, M., & Mandich, A. (2006). School-based occupational therapy for children with fine motor difficulties: Evaluating functional outcomes and fidelity of services. *Physical & Occupational Therapy in Pediatrics*, *26*, 89-110.

Bellg, A.J., Borrelli, B., Resnick, B., Hecht, J., Sharp Minicucci, D., Ory, M. et al. (2004). Enhancing treatment fidelity in health behaviour change studies: Best practices and recommendations from the NIH Behaviour Change Consortium. *Health Psychology, 23,* 443-451.

Billingsley, F., White, O.R., & Munson, R. (1980). Procedural reliability: A rationale and an example. *Behavioral Assessment,* 2, 229-241.

Bond, G. R., Evans, L., Salyers, M., Williams, J., & Hea-Won, K. (2000). Measurement of fidelity in psychiatric rehabilitation. *Mental Health Services Research, 2,* 75-87.

Boutron, I., Moher, D., Altman, D.G., Schultz, K.F., & Ravaud, P. (2008). Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment: Explanation and Elaboration. *Annals of Internal Medicine, 148,* 295-306.

Breitenstein, S.M., Gross, D., Garvey, C.A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation Fidelity in Community-Based Interventions. *Research in Nursing & Health, 33,* 164-173.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science 2*, 40 doi:10.1186/1748-5908-2-40

Carroll, K.M., Nich, C., Sifry, R.L., Nuro, K.F., Frankforter, T.L., Ball, S.A. (2000). A general system for evaluating therapist adherence and competence in psycho-therapy research in the addictions. *Drug and Alcohol Dependence, 57,* 225-238.

Chambless, D.L., & Hollon, S.D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66,* 7-18.

Cited Reference Search #1: Cited Reference Index for Waskow, IE (1984) from Psychotherapy Research. Retrieved on April 16, 2012, from the Web of Science database.

Cited Reference Search #2: Cited Reference Index for Lichstein KL et al. (1994) from Advances in Behaviour Research and Therapy. Retrieved on April 16, 2012, from the Web of Science database.

Cited Reference Search #3: Cited Reference Index for Carroll C et al. (2007) from Implementation Science. Retrieved on April 16, 2012, from the Web of Science database.

Coster, W., Tickle-Degnen, L., and Armenta, L. (1995). Therapist-child interaction during sensory integration treatment: Development and testing of a research tool. The Occupational Therapy Journal of Research, 15, 1, 17-35.

Craig, P., Dieppe, P., Macintyre, S., Mitchie, S., Nazareth, S., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal, 337,* 979-983.

Dane, A., & Schneider, B. H. (1998). Integrity in primary prevention programs: Are implementation effects out of control? *Clinical Psychology Review, 18,* 23-45.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., Donaldson, A., & Varley, J. (2010). Randomized, Controlled Trial of an Intervention for Toddlers With Autism: The Early Start Denver Model. *Pediatrics, 125,* e17-e25.

Dijkers, M., Kroop, G.C., Esper, R.M., Yavuzer, G., Cullen, N., Bakdalieh, Y. (2002). Quality of intervention research reporting in medical rehabilitation journals. *Amer J Phys Med, 81,* 21-33.

Di Rezze, B., Law, M., Gorter, J.W., Eva, K., and Pollock, N., (In Press). A narrative review of generic intervention fidelity measures. *Physical and Occupational Therapy in Pediatrics.*

Dunn, W. (2011). Best practice occupational therapy for children and families in community settings. Thorofare, NJ: Slack

Dunst, C.J. (2009). Implications of Evidence-Based Practices for Personnel Preparation Development in Early Childhood Intervention. *Infants and Young Children, 22,* 44-53.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. (2003) A Review of Research on Fidelity of Implementation: Implications for Drug Abuse Prevention in School Settings. *Health Education Research 18,* 237-256.

Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28,* 1167-1187.

Eva, K. (2003). On the generality of specificity. *Medical Education, 37,* 587-588.

Eysenck, H. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16,* 319-324.

Green, J. (2006). Annotation: The therapeutic alliance – a significant but neglected variable in child mental health treatment studies. *Journal of Child Psychology and Psychiatry, 47,* 425 - 435.

Gresham, F.M. (1997). Treatment integrity in single-subject research. In R. D. Franklin, D.B. Allison, & B.S. Gorman (Eds.), *Design and analysis of single-case research* (pp.93-117). Mahwah, NJ: Erlbaum.

Hart, T. (2009). Treatment definition in complex rehabilitation interventions. *Neuropsychological Rehabilitation, 19,* 824-840.

Hennessey, M.L., & Rumrill, P.D. (2003). Treatment fidelity in rehabilitation research. *Journal of Vocational Rehabilitation, 19,* 123-126.

Hoagwood, K., Hibbs, E., Brent, D., and Jensen, P. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 683-687.

Hogue, A., Liddle, H.A., Singer, A., & Leckrone, J. (2005). Intervention fidelity in family-based prevention counseling for adolescent problem behaviors. *Journal of Community Psychology, 33,* 2, 191-211.

Hollon, S.D. (1984). Final Report: System for rating psychotherapy audiotapes. Rockville, MD: Department of Health and Human Services.

Institute of Medicine (2001). *Improving the quality of long-term care.* National Academy Press, Washington, DC, USA.

Judge Santacroce, S.J., Maccarelli, L.M. & Grey, M. (2004). Intervention fidelity. *Nursing Research, 53,* 63-66.

Kazdin, A.E. (1980). *Research design in clinical psychology.* New York: Harper & Row.

Kendall, P.C., & Beidas, R.S. (2007). Smoothing the trail for dissemination of evidence-based practices for youth: Flexibility within fidelity. *Professional Psychology: Research and Practice, 38,* 13-20.

Ketelaar, M., Kruijsen, A., Verschuren, O., Jongmans, M., Gorter, JW, Verheijden, J., Reinders-Messelink, H., Lindeman, E. (2010). LEARN 2 MOVE 2-3: a randomized controlled trial on the efficacy of child-focused intervention and context-focused intervention in preschool children with cerebral palsy. *BMC Pediatrics, 10*:80. doi:10.1186/1471-2431-10-80

King, S., Rosenbaum, P., & King, G. (1995). The Measure of Processes of Care: A means to assess family-centred behaviours of health care providers. Hamilton, ON: McMaster University, Neurodevelopmental Clinical Research Unit.

Kraemer, H., Wilson, T., Fairburn, C., & Agras, W.S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59,* 877-883.

Lambert, M. J., &, Ogles, B. M. (2004). The Efficacy and Effectiveness of Psychotherapy. In M.J. Lambert, *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (5th ed.) pp. 139–193. New York: John Wiley & Sons.

Lane, S., and Bundy, A. (2012). Kids Can Be Kids: A Childhood Occupations Approach. Philadelphia, PA: F.A. Davis Company.

Law, M., Darrah, J., Pollock, N., Wilson, B., Russell, D., Walter, S., Rosenbaum, P., Galuppi, B. (2011). Focus on function: a cluster, randomized controlled trial

comparing child- versus context-focused intervention for young children with cerebral palsy. *Developmental Medicine and Child Neurology,* 53, 621-629.

Law, M., Missiuna, C., Pollock, N., and Stewart, D. (2001). Foundations for Occupational Therapy Practice with Children. In J. Case-Smith (Ed.), *Occupational Therapy for Children.* St. Louis, MI: Mosby, Inc.

Leventhal, H.E., & Friedman, M.A. (2004). Does Establishing Fidelity of Treatment Help Us to Understand Treatment Efficacy? *Health Psychology, 23,* 452-456.

Lichstein, K.L., Riedel, B., & Grieve, R. (1994). Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research & Therapy, 16,* 1-29.

Markowitz, J.C., Spielman, L.A., Scarvalone, P.A., & Perry, S.W. (2000). Psychotherapy adherence of therapists treatment HIV-Positive patients with depressive symptoms. *Journal of Psychotherapy Practice Research, 9,* 2, 75-80.

Martino, S., Ball, S.A., Nich, C., Frankforter, T.L., & Carroll, K.M. (2008). Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence, 96,* 37-48.

Mattingly, C., and Fleming, M. (1994). *Clinical Reasoning: Forms of Inquiry in a Therapeutic Practice.* Philadelphia, PA: F.A. Davis Company.

Medical Research Council (2000). A framework for development and evaluation of RCTs for complex interventions to improve health 2000. [ONLINE] Available at http://www.mrc.ac.uk/complex-interventionsguicance. Accessed on April 16, 2012.

Michaud, L. (2004). Prescribing therapy services for children with motor disabilities. *American Academy of Pediatrics, 113,* 1836-1838.

Miller, L. J., Schoen, S. A., James, K., & Schaaf, R. C. (2007). Lessons learned: A pilot study on occupational therapy effectiveness for children with sensory modulation disorder. *American Journal of Occupational Therapy, 61,* 161–169.

Moll, S., Huff, J., & Detwiler, L. (2003). Supported employment: Evidence for a best practice model in psychosocial rehabilitation, *70,* 298-310.

Moncher, F.J., & Prinz, R.J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11,* 247-266.

Morgenstern, J., & McKay, J.R. (2007). Rethinking the paradigms that inform behavioral treatment research for substance use disorders. *Addiction, 102,* 1377-1389.

Nelson, D.L., & Mathiowetz, V. (2004). Randomized controlled trials to Investigate occupational therapy research questions. *American Journal of Occupational Therapy, 58,* 24-34.

Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E. & Galuppi, B. (1997). Development and validation of a gross motor function classification system for children with cerebral palsy. *Developmental Medicine and Child Neurology*, *39*, 214–223.

Parham, D., Roley, S., May-Benson, T., Koomar, J., Brett-Green, B., Burke, J.P., Cohn, E., Mailloux, Z., Miller, L.J., Schaaf, R., (2011). Development of a fidelity measure for research on the effectiveness of the Ayres Sensory Integration intervention. *American Journal of Occupational Therapy, 65,* 133-142

Quay, H.C. (1977). The three faces of evaluation: What can be expected to work. *Criminal Justice and Behavior, 4*, 341-354.

Rodger, S., & Keen, D. (2010). Child- and Family-Centred Service Provision. In S. Rodger (Ed.), *Occupation-centred Practice with children* (pp. 45-69). West Sussex, UK: John Wiley & Sons, Inc.

Rose, B., Holmbeck, G., Coakley, R., & Franks, E. (2004). Mediator and moderator effects in developmental and behavioral pediatric research. *Developmental and Behavioral Pediatrics, 25*, 58-67.

Rosenbaum, P., King, S., Law, M., King, G., and Evans, J. (1998). Family-Centred Service: A conceptual framework and research review. *Physical & Occupational Therapy in Pediatrics, 18*, 1, 1-20.

Rounsaville, B.J., & Carroll, K.M. (2001). A stage model of behavioral therapies research: Getting started and moving on from stage 1. *Clinical Psychol Sci Prac, 8*, 133-142.

Schlosser, R.W. (2002). On the importance of being earnest about treatment integrity. *Augmentative and Alternative Communication, 18*, 36-44.

Sechrest, L., West, S.G., Philips, M.A., Redner, R., & Yeaton, W. (1979). Some neglected problems in evaluation research: Strength and integrity of treatments. In L. Sechrest, S.G. West, M.A. Philips, R. Redner, & W. Yeaton (Eds.), *Evaluation studies review annual* (Vol.4, pp. 15-35). Beverly Hills: Sage.

Smith, S.W., Daunic, A.P., Taylor, G.G. (2007). Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. *Education and Treatment of Children, 30,* 4, 121-134.

Stein, K.F., Sargent, J.T., & Nicholas, R. (2007). Intervention research: Establishing fidelity of the independent variable in nursing clinical trials. *Nursing Research, 56,* 54-62.

Streiner, D.L. & Norman, G.R. (2006). *Health Measurement Scales: A practical guide to their development and use 3rd Ed.* New York: Oxford University Press.

Waltz, J., Addis, M.E., Koerner, K. & Jacobson, N.S. (1993). Testing the integrity of psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61,* 620-630.

Waskow, I.E. (1984). Specification of the technique variable in the NIMH Treatment of Depression Collaborative Research Program. In: Williams, J.B.W., Spitzer, R.L. (Eds.), Psychotherapy Research: Where are we, and where should we go? New York: Guilford.

Whyte, J. (2008). A grand unified theory of rehabilitation (We Wish!). The 57th John Stanley Coulter Memorial Lecture. *Archives of Physical Medicine and Rehabilitation, 89,* 203-209.

Whyte, J. (2003). Clinical trials in rehabilitation: What are the obstacles? *American Journal of Physical Medicine and Rehabilitation, 82,* 10(Suppl), S16-S21.

Whyte, J. & Hart, T. (2003). It's more than a black box; It's a Russian doll: Defining rehabilitation treatments. *American Journal of Physical Medicine and Rehabilitation*, 82, 639-652.

Yeaton, W.H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity and effectiveness. *Journal of Consulting and Clinical Psychology, 49,* 156-167.

APPENDICES

Appendix A: Glossary of Terms

Child Therapy – A traditional intervention approach in paediatric rehabilitation that involves therapeutic strategies to remediate or make changes to the child`s ability to help them meet their therapy goal.

Child Therapy Domain – Refers to the body of items on the PROF that reflect the strategies of Child therapy.

Context Therapy – A novel intervention approach in paediatric rehabilitation that involves an approach to making changes to the child`s environment or activity to help them meet their therapy goal.

Context Therapy Domain – Refers to the body of items in the PROF that reflects the strategies of Context therapy.

General Therapy Behaviours Domain – Refers to the group of observable behaviours involving the child, parent and therapist within either of the interventions delivered in this study.

Intervention non-specific behaviour or general therapy behaviour – Any behaviours of the therapist, client (child and/or parent) or interaction between the parties that can be observed in any therapy session (i.e., non-specific to the intervention-type).

Intervention specific behaviour – Any therapist behaviours that can be observed in an intervention session for a particular intervention approach.

Generic Fidelity Measure – a measure that examines unique behaviours relevant to each intervention-type, as well as non-specific intervention behaviours

Appendix B:

The Paediatric Rehabilitation Observational Measure of Fidelity Manual

Measure of Behaviours in Paediatric Intervention:

Rating Manual

Briano Di Rezze

McMaster University

N.B.: to ensure that the raters were naïve to the purpose and hypotheses in the development of the PROF, the measure was called the Measure of Behaviours in Paediatric Intervention (MBPI) during the training and psychometric testing.

Table of Contents

This manual is designed to provide raters with a guide for scoring videotapes of paediatric rehabilitation intervention sessions using the Measure of Behaviours in Paediatric Intervention (MBPI). The measure examines behaviours of those involved in a paediatric rehabilitation intervention session, including the therapist, child and parent (if in attendance). The manual is a reference document when rating videotaped intervention sessions. The manual is divided into the following sections: introduction to the measure; general guidelines for rating the measure; using the rating scales; item descriptions; and rating materials.

Introduction to the measure

Purpose: The Measure for Behaviours in Paediatric Intervention (MBPI) is used to rate observable behaviours of participants in a paediatric rehabilitation session on videotape. The videos will involve occupational therapists and physiotherapists conducting interventions with children with cerebral palsy. Ratings of the videos will help to collect information about behaviours that are evident in a therapy session.

Overview of the measure: The MBPI has **four sections** of items that include both behaviours that are intervention-specific and general behaviours observed in any intervention session.

- The **first section** consists of behaviours focusing on intervention content potentially demonstrated by the **therapist**.
- The **second section** has items focused on general behaviours within the intervention session potentially demonstrated by the **therapist**. These behaviours can be applicable to any intervention session across paediatric clients.
- The **third section** examines potential general behaviours of the **child and/or the parent** demonstrated within an intervention session.
- The **fourth section** examines the interaction behaviours demonstrated within the intervention session between the **therapist and the child or parent**.

In each of these sections, the list of behaviours across participants (i.e., therapist, child/parent, or interaction combinations) are rated based on the **frequency** of each item behaviour on video. For the therapist behaviour items, a second rating is also expected to rate the **quality** of the observed behaviour.

General Rating Guidelines for Videotaped Intervention Sessions

1. The MBPI is to be used with a videotaped intervention session. Prior to observing each videotape, **familiarize yourself with all sections and 28 item behaviours** on the MBPI. Always have the scoring manual present to ensure the examination of the correct behaviours. Refer to the manual whenever there is any doubt or uncertainty.

2. Review the entire videotaped intervention session before rating the items. The sessions will vary in length from 20-50 minutes. All items are to be rated for each videotape **(do not leave any items blank)**. Item ratings are based on what is **observed or heard** within the intervention session. To determine ratings, take notes on the worksheet or in the space provided on the measure while observing the video. Use these notes to help jog your memory regarding the frequency of observed behaviour as well as the quality of therapist behaviour. Using the scale reference (page 6 of the measure) clearly write the score in the textbox next to each item based on the adjective that best fits what you observed on video

3. All items on the MBPI are to be rated on a scale of Frequency. When rating the items, base the scores solely on the frequency to which the behaviour occurs in the videotaped intervention session. For therapist behaviour items (items 1-16), a second scale is present to rate the quality of the observed behaviours, based on your clinical knowledge of intervention in paediatric rehabilitation. These scales examine different aspects of the behaviour, so do not assume that higher frequency automatically equals a high quality rating. **It is possible to observe a behaviour once, performed at high quality.**

4. Rate only the behaviours that you **see or hear on video** in the intervention session. Do not assume what a therapist might be thinking; behavioural evidence is essential.

5. When observing the videotaped sessions, it is normal to think about how you might do things similarly or differently in an intervention session. These judgements **should not influence your ratings** on the scales provided. Prior to rating therapist performance on an item, **examine and eliminate any personal judgements** of the therapy style.

6. As in any given intervention session, all item behaviours may not be observed in every videotaped session and therapist performance may focus on some behaviours more frequently than others. Similarly, the observance of child and/or parent

behaviours may also vary from session to session. **Just because a behaviour was not observed, it does not mean that the therapist did a poor job in providing intervention.**

7. Avoid letting earlier item ratings influence the rating of other items for the same person.

8. All videotapes and rating scores are confidential materials. Observe videos and rate them in a location where colleagues, family members and/or friends cannot see or hear the sessions. You should handle videos like medical charts and make sure to not leave them unattended. Do not discuss the names of these individuals and/or session content with anyone other than the study investigator.

Scales to Rate Items on the MBPI

On the MBPI, there are two scales to be used to rate each item: **(1) Frequency** of the observed behaviour; **and** (2) **Quality** of the observed behaviours**.**

(1) Frequency of Observed Behaviours Scale
The main scale in the MBPI examines the frequency of the behaviours demonstrated by the therapist, child or parent, or interaction between them in an intervention session. **Frequency** is the number of instances in which item behaviours are observed throughout the course of the intervention session. A score is based on the content of the entire videotaped intervention by the person expected to perform specific behaviours (e.g., therapist, child, parent or interactions).
The MBPI should reflect the raters' approximate estimate of the frequency observed for each item, see the scale below.

Note: The frequency of child/parent behaviours as well as behaviours whereby the child interacts with the therapist are to be scored while taking into consideration the **perceived abilities of the child and the expectations of performance for his/her approximate age**. Make sure that you are rating the identified person or interaction between people for each item and rate **N/A if the person is not present or heard** on video.

Frequency Scale

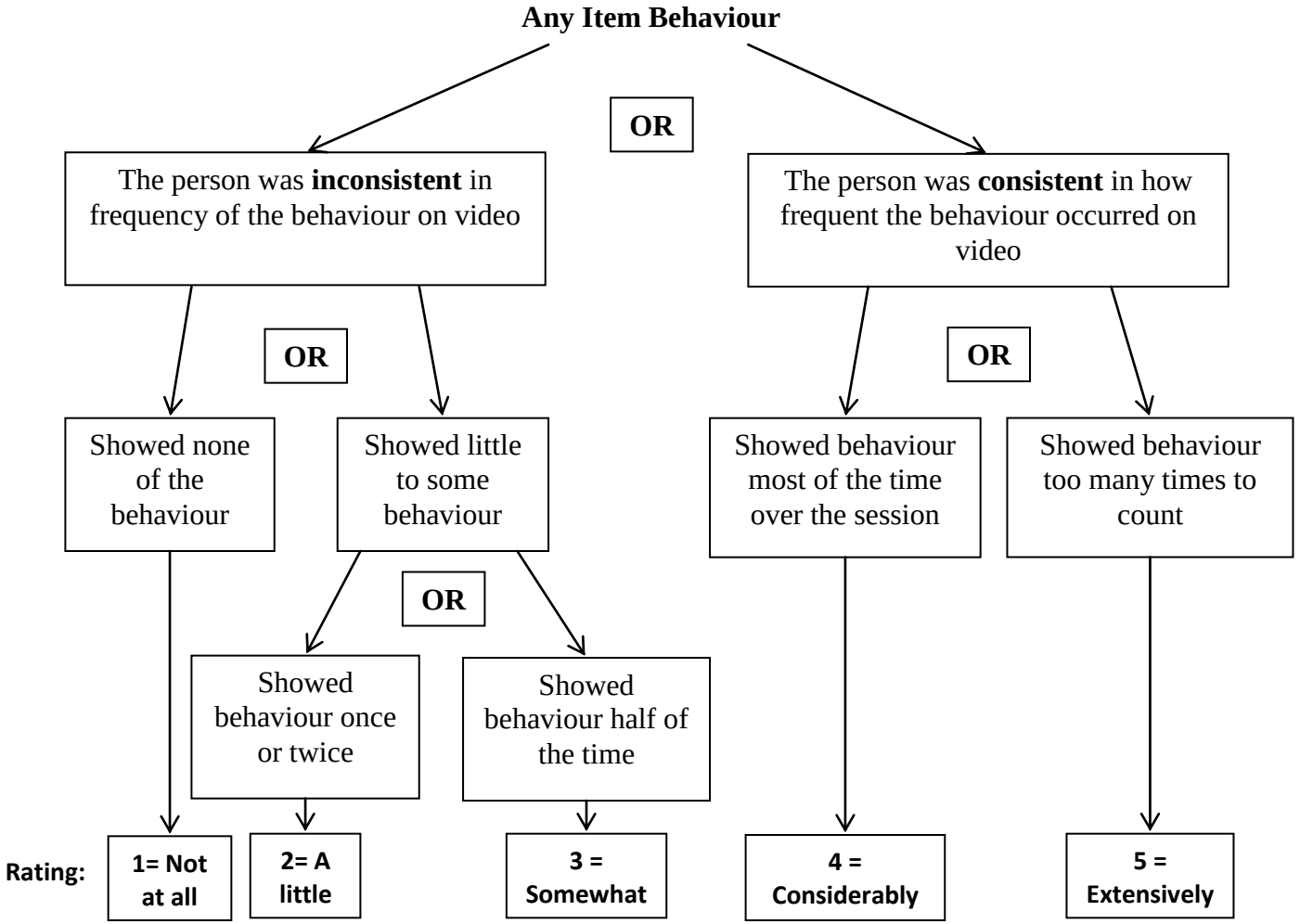
Score:	N/A	1	2	3	4	5
Description:	Not present	Not at all	A little	Somewhat	Considerably	Extensively

The Likert rating scale contains adjectives to identify what was observed. The scale is

designed to assess the frequency to which the participants engage in the specific intervention session behaviours. Descriptions of the meaning of each adjective is provided below.

Description of rating options for Frequency of Behaviours:

N/A = Not Applicable**,** if the person to be rated **was not present in the video**.
1 = Not at all. The person to be rated **was present or heard** and the item behaviour never explicitly occurred.
2 = A little. The item behaviour occurred once or twice over the course of the session.
3 = Somewhat. The item behaviour occurred **about half of the time** over the course of the session.
4 = Considerably. The item behaviour occurred **most of the time** over the course of the session.
5 = Extensively. The item behaviour was **integrated within the entire** intervention session, so much so that it became difficult to count frequency.
To help with **rating the frequency of behaviours** for each item, raters could use the **flowchart below** to facilitate the decision-making process:


Any Item Behaviour
OR
The person was **inconsistent** in frequency of the behaviour on video
The person was **consistent** in how frequent the behaviour occurred on video
OR
OR
Showed none of the behaviour
Showed little to some behaviour
Showed behaviour most of the time over the session
Showed behaviour too many times to count
OR
Showed behaviour once or twice
Showed behaviour half of the time
Rating:
1= Not at all
2= A little
3 = Somewhat
4 = Considerably
5 = Extensively

Avoid attempts to estimate the quality with which interventions are executed or the appropriateness of the intervention for the given situation. Quality of the item behaviour will only be scored on items focusing on therapist behaviours in a second scale "Quality of Behaviours" (See section below).

(2) Quality of therapist behaviours Scale

For each item, characteristics of quality to consider for therapist's behaviours are **observed expertise, appropriate timing of behaviours, clarity of language (appropriate to the child and/or parent), and appropriate actions for a particular situation (or response to the client's needs)**. Ensure that the emphasis is on these specific characteristics, **NOT** ones related to other items, such as, warmth, fun, rapport. If the rater has difficulty scoring therapist performance items, first think about the best and worst-case scenarios relevant to the item, then consider how the quality of observed behaviours from the video compares to clinical practice experience.

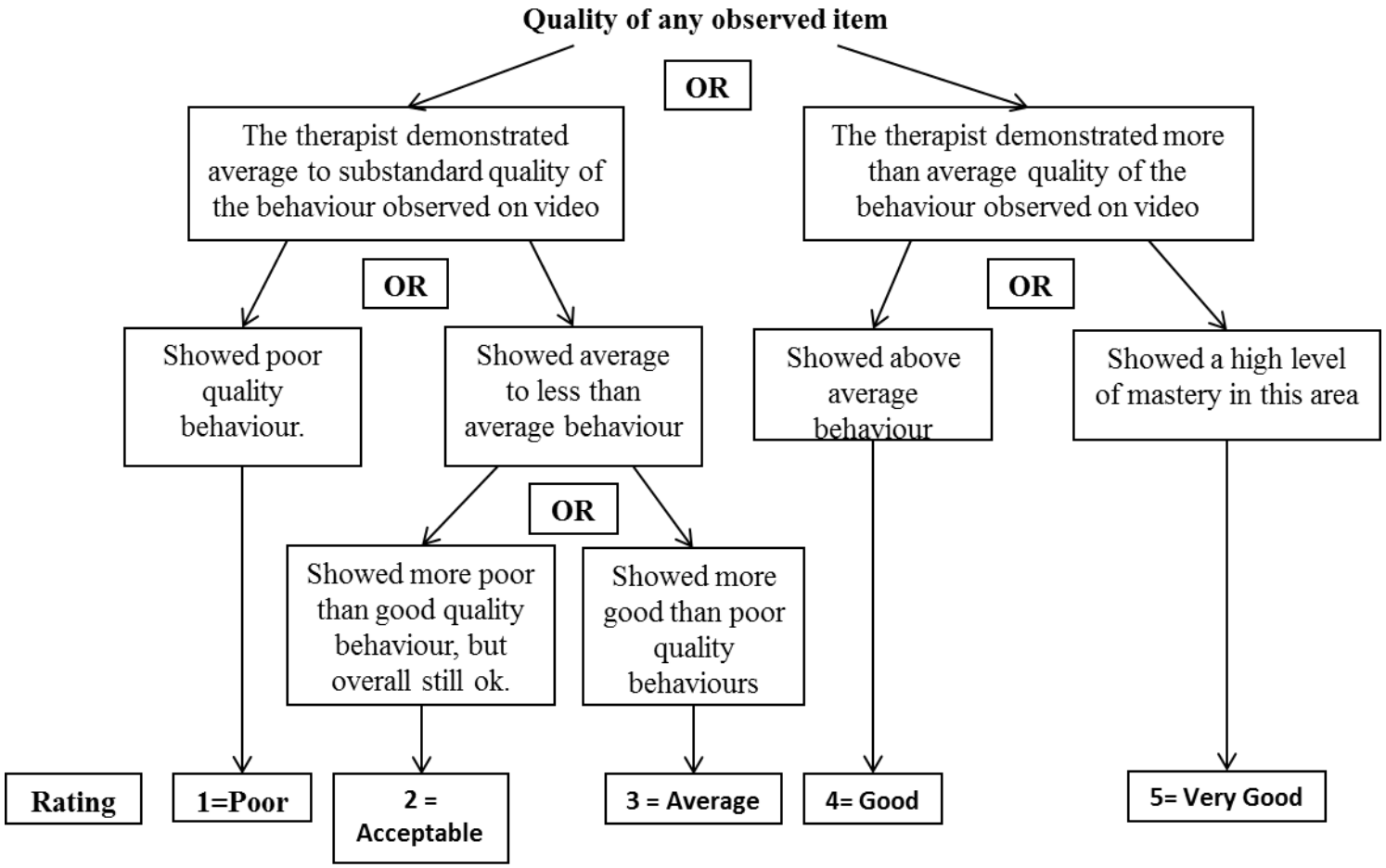
Quality of Therapist Behaviour Scale

Score:	N/ A	1	2	3	4	5
Description:	Freq. rating =1	Poor	Acceptable	Average	Good	Very Good

As evident in the scale above, the adjectives for quality of behaviour differ from those of the frequency of behaviours scale. Descriptions of the adjectives are provided below.
N/A = Not Applicable**,** since the **therapist behaviour** was not observed (i.e., rated "1" on Frequency scale).

1 = Poor. Therapist handled the described item behaviour poorly (e.g., showing clear lack of expertise, understanding, competence, commitment, inappropriate timing, and unclear language).
2 = Acceptable. Therapist behaviour for an item was observed as '**good enough**,' but in a less than 'average' manner based on the dominance of less critical factors (e.g., unclear language).
3 = Average. Therapist behaviour for an item was observed as 'average', where the therapist showed **mostly good than poor quality** behaviours.
4 = Good. Therapist behaviour was observed to be **better than 'average'** within an intervention session at a level that demonstrated skill and expertise in performing the item behaviour.
5 = Very good. Therapist behaviour was demonstrated at a **high level of mastery** in this area.

To help with **rating the quality of the therapist's behaviour** for each item, the raters could use the **flow chart below** to facilitate the decision-making process:


Quality of any observed item
OR
The therapist demonstrated average to substandard quality of the behaviour observed on video
The therapist demonstrated more than average quality of the behaviour observed on video
OR
OR
Showed poor quality behaviour.
Showed average to less than average behaviour
Showed above average behaviour
Showed a high level of mastery in this area
OR
Showed more poor than good quality behaviour, but overall still ok.
Showed more good than poor quality behaviours
Rating
1=Poor
2 = Acceptable
3 = Average
4= Good
5= Very Good

Overall, **remind yourself** that for **Frequency** ratings you are rating the behaviours that are observed. For therapist behaviours, **Quality** ratings will also be scored for each therapist behaviour item, whereby the focus is on the using your clinical judgement of how well the therapist item behaviour was implemented (e.g., appropriateness, timing of it, etc.). Items related to **therapist behaviours** that are not observed or heard will be scored a "1" on the Frequency scale and automatically "Not Applicable (N/A)" on the Quality of behaviour scale.

Definitions of key terms

Activity: (noun) Over the course of an intervention session, an activity could be observed as an event that is purposeful for the child to participate in. Such events may have a start and end (e.g., a game of snakes and ladders). Furthermore, events may take the form of practicing a repetitive task with or without clear guidelines to begin and end (e.g., strengthening exercises). This also includes activities that the child may engage in as a passive recipient (e.g., warm-up stretches).

Abilities: "General traits which are a product of genetic make-up and learning, e.g., spatial orientation is necessary for success at reading a road map and finding your way through a familiar room" (Christiansen & Baum, 1991, p.23).

Skills: "Pertains to the level of proficiency in a specific task…skill in complex tasks can be explained by the presence of various underlying general abilities, e.g., professional driver has more driving skills than the average person" (Christiansen & Baum, 1991, p.23). Skills are essentially, the demonstration of a combination of abilities for a complex task.

Environment: Social and physical conditions external to the individual or factors which have the potential to influence the individual's performance in a particular context (e.g., equipment to assist interaction within the environment) (Christiansen & Baum, 1991).

Remediation: Actions by the therapist that attempts to change the psychological, cognitive, physiological, and neurobehavioural abilities of an individual (Christiansen & Baum, 2005).

Item Descriptions

The MBPI is comprised of 28 items. In table 1, a definition of each item is provided. Where necessary **distinctions are provided for an item** that may seem similar on first glance to another item but are in fact different. These distinctions are not intended to be an exhaustive list for each item, and raters may need to make some unanticipated distinctions.

Table 1: Item Descriptions on the MBPI

THERAPIST BEHAVIOURS (INTERVENTION-SPECIFIC)
1. Did the THERAPIST focus on **remediating** the CHILD'S movements or skills? **Item Definition:** This item measures how the therapist focused the intervention on **changing the abilities or skills of the child** within the intervention session. This item may take the form of the therapist guiding the child within an activity, through verbal or hand over hand instruction, or specific strategies (i.e., stretching to increase ROM or sensory based input) with the intent of changing the child's baseline abilities.
2. Did the THERAPIST focus on the CHILD **completing activities in a specific way** (e.g., follow a specific sequence of steps)? **Item Definition:** This item measures whether the intervention focus aimed to systematically develop skills and build on them to work towards a bigger skill (e.g., proximal motor abilities addressed before distal motor abilities). Here the therapist is requesting the child to perform a specific sequence related to the therapy, rather than giving the child instructions to do the activity in whatever way works for them. **Distinctions from other items:** Different from item #6 whereby the therapist does not focus on how the activity is completed but by any means enables the child to complete the activity (i.e., in #6 the therapist just wants the child to "get 'er done").
3. Did the THERAPIST focus on the **quality** of the CHILD'S movement and/or skills? **Item Definition:** This item measures how the therapists' intervention took into consideration how smoothly the child demonstrated his/her abilities or skills. Potential examples of what the therapist could request to facilitate better proficiency in the child's performance include having the child slow down his/her performance or provide direct feedback on body position (e.g., straighten arm, or move in a particular way).
4. Did the THERAPIST facilitate **repetition in the practice** of movements or motor skills addressed in the intervention session? **Item Definition:** This item measures how the therapist uses the intervention as an

opportunity to teach and **practice skills** introduced in the session to change the abilities or skills of the child. Movements or motor skills could also focus on positioning of the child.

Distinctions from other items: Different from item #7 that emphasizes the focus on the child **practicing an activity** through completion, rather than practicing for repetition in this item.

5. Did the THERAPIST focus on **changing** the **environmental constraints** that influence the CHILD'S activity performance (e.g., environmental barriers within an activity)?

Item Definition: This item measures how the therapist modifies elements of the environment within the intervention session to enable the child to participate in the activities. Elements of the environment also include providing some assistance such as introducing equipment (e.g., electronic switches), or changing the working surface or environment. Physical assistance is not included, e.g., hand over hand strategies. Note that social nuances that could be performed by any adult outside of the context of a therapy session should not count as an observation for this item. For example, a therapist moving the table closer to a child is something that could be expected by any adult interacting with a child and is not an intervention-specific strategy exclusively used by a therapist.

6. Did the THERAPIST focus on **adapting an activity** to enable the CHILD to complete the activity?

Item definition: This item measures whether the therapist makes changes to the activity in a way that it will **regularly be performed** by the child within and/or outside of the intervention session. The important characteristic of this item is that the child will be expected to perform this different or adapted activity in order to make it easier for him/her to complete the activity (e.g., changing the activity sequence or simplifying the steps in how the activity is done).

Distinctions from other items: This item differs from item #12 (grading the level of difficulty in steps of an activity) in that here an activity is adapted with the intention of enabling the child to participate in the activity successfully (i.e., the child can be expected to use this adapted approach to perform the activity). On the other hand, **grading the activity** would involve the therapist making adjustments in order to work on **specific steps** within an activity (e.g., holding a container of toys closer to the child to enable the child to focus on their performance within the activity).

7. Did the THERAPIST provide opportunities for the CHILD to **practice activities for the purpose of completion** (i.e., not concerned with how the activity was done)?

Item Definition: This item measures how much the therapist emphasized practicing activities to their completion (e.g., functional goal), rather than focus on the specific component skills of the child in order to complete the activity.

Distinctions from other items: This item differs from #6 in that this item focuses on whether activities are practiced (more than once) to completion. Whereas, item #6 focuses only on whether the activity was adapted for the child to do to facilitate completion, which only needs to be observed (not repeated).

8. Did the THERAPIST focus on the **most efficient way that best suits the CHILD** to perform the activity?

Item Definition: This item measures how much the therapist focused on completing the activity in the intervention session using the quickest method possible that is easiest on the child, directed by the child, or is the preferred solution of the child.

Distinctions from other items: In contrast to item #7, whereby the therapist provides opportunities to practice activities to completion, this item focuses on **how** activity was completed.

THERAPIST BEHAVIOURS (GENERAL PRACTICE)

9. Did the THERAPIST **assess** the CHILD (formally or informally)?

Item Definition: This item is intended to measure the extent to which the therapist assessed the child's abilities within the intervention session. The assessment does not have to occur only at one time within the session. An assessment may be as simple as requesting the child to perform certain movements and/or activities. It may also be conducted formally or informally to examine abilities and skills.

10. Did the therapist **review therapy goals and/or activities** for the CHILD (and/or Parent) **to practice outside of the intervention** sessions?

Item Definition: Overall, the therapists' actions are meant to examine goal progress and/or reinforce what was done in the intervention session to facilitate goal achievement. Some behaviours involve how the therapist facilitates the dialogue around the child's goals to suit the needs of the **child and family** with respect to intervention plan. Emphasis is on the discussion of the goals to determine whether they are still appropriate or what components of the goals need to be further addressed. Note that the mere mention of goals covered in a session is not enough to result in an observed behaviour.

Other behaviours appropriate for this item measures how the therapist provides **explicit** instruction and/or information to parent regarding the strategies or approaches used in the intervention session **in order to integrate** therapy into regular routine in family's home life (e.g., therapist models or provides instruction on technique(s) used in session for parents to try at home).
11. Did the THERAPIST **demonstrate, instruct and/or teach** the CHILD skills or strategies? **Item Definition:** This item measures how the therapist spent time in the session teaching the child (or parent, when the child could not perform activities independently) by him/herself through practice, didactic teaching, or activity exploration. This behaviour is in contrast to the child playing freely for the entire time.
12. Did the THERAPIST **grade** the level of **difficulty of the steps** to participate in an activity? **Item Definition:** This item measures how the therapist mediates between the child's abilities and the activity to facilitate the appropriate challenge for the child to **practice part(s) of an activity.** These parts or steps to the activity would be expected to increase in difficulty over time (e.g., hand positioning to provide different levels of physical support over time). For this item, grading steps of an activity could also mean making adjustments to how long the steps of an activity are practiced**,** such as terminating an activity early to enable the child to have a positive experience. **Distinctions from other items: See item #7 for distinction.** Item #6 focuses on adapting the activity with the emphasis of performing and completing the activity that has been modified in a way that it will regularly be performed.
13. Did the THERAPIST offer the CHILD **activities** that were **engaging or fun**? **Item Definition:** This item measures to what extent the therapist provides activities that encourages the child to engage in pleasurable activities. This includes the therapists' consideration of the child's preferences. An example of an observable behaviour to determine if the child is engaged or having fun could be based on whether the child is observed as happy or going along with the activity.
14. Did the THERAPIST **praise or communicate confidence** in the CHILD'S effort or performance? **Item Definition:** This item measures how supportive the therapist is of the child`s efforts and achievements in an intervention session. The item looks closer at the therapist`s

ability to create a positive learning environment by celebrating or rewarding achievements in session.
15. Did the THERAPIST **demonstrate warmth towards** the CHILD? **Item Definition:** Warmth is understood as unconditional positive regard and is defined as "the therapist communicating to his/her client a deep genuine caring for his/her as a person with human potentialities, a caring uncontaminated by evaluations of his thoughts, feelings, or behaviours." This communication need not be explicit, but the therapist's caring should be evident by his/her behaviour (e.g., touch, tone, or facial expressions). **Distinctions from other items:** Different from item #25 (therapist-child rapport). For this item, it is possible for the therapist to be warm and caring and yet not get along with the child. Conversely, it is possible for the therapist to not demonstrate warmth or caring for the child and yet develop strong rapport. This item is also different from item #14 (praise or communicating confidence) in that a therapist can use the appropriate phrases to praise the child, but may not demonstrate the characteristics of being warm as indicated in the description above.
16. Did the THERAPIST alter his/her behaviour (OR show flexibility) **to adapt to the needs, behaviours, or interests** of the CHILD? **Item Definition:** This item measures to what extent the therapist demonstrates his/her ability to address the needs that arise in a therapy session, whether or not it meets their agenda. E.g., Addressing family/child needs, goals, or activities based on psychosocial, situational or disability-focused changes prior to the intervention session.
CHILD OR PARENT BEHAVIOURS
17. Did the CHILD **actively participate** in the intervention session? **Item Definition:** This item measures the child's interest and/or involvement in session activities**.** These behaviours may manifest themselves as amount of child participation, attitude towards the therapist or the therapy session, child affect, body language, and level of interest in the discussion and activities that occur in session. Try not to consider whether or not the child "likes" the therapist or the therapy, but whether or not he or she is receptive to and/or invested in the session.
18. Did the CHILD **show comfort** with the therapist? **Item Definition:** This item measures the extent to which the child has confidence (or trust) in the therapist within the intervention session.

19. Did the CHILD **communicate his/her preferences**?

Item Definition: This item measures the extent to which the child vocalizes his/her needs, wishes, and refusals, as well as demonstrates his/her pleasure, excitement and displeasure within the context of the intervention session. If the child is unable to do so independently, the parent may be vocal about what is in the best interest of their child.

20. Did the CHILD **show discomfort due to a personal factor** (e.g., physical pain, sleepiness, emotional lability)?

Item Definition: This item measures the extent to which the child demonstrates some difficulty (or discomfort) in the session, most likely evident by the mood of the child (i.e., "under the weather"). Other observations of the child that show discomfort could include sleepiness or pain.

Distinction from other items: This item is distinct from item #21 in that this item focuses on discomfort that is explainable due to a factor affecting the child (within the child). On the other hand, item #18 focuses on the child's comfort within the session based on factor(s) involving the therapist. This item also differs from item #19, whereby the child could be communicating his/her displeasure within session for reasons that seem to be typical defiance behaviour by a child.

21. Did the CHILD (or PARENT where appropriate) **practice strategies with the therapist**?

Item Definition: This item measures the extent to which the child actively practices strategies taught within the intervention session. If the child is not able to physically or cognitively be independent in performing strategies taught, then the parent may facilitate the practice of strategies for the child (i.e., child is a passive recipient).

Distinction from other items: Distinct from Item #20 that focuses on the parent participating in the intervention session because the techniques require the parent to facilitate the child's performance. Item #21 **may** require the parent to assist the child with practicing the strategies if the child is unable to do so independently and the therapist requires assistance. If the child is able to do so independently, then this item does not involve the parent and is clearly distinct from item #20.

22. Did the PARENT **participate** in the intervention session?

Item Definition: This item measures the extent to which the parent independently engages within session and uses strategies to address the child's goal.

23. Did the PARENT **appear to be comfortable** with the therapist? **Item Definition:** This item measures the extent to which the parent is at ease with intervention session and intervention discussions and/or strategies do not appear to be a source of stress for them.
24. Did the PARENT **communicate his/her preference(s)**? **Item Definition:** This item measures the extent to which the parents are involved in the session by communicating their thoughts and concerns with the therapist.
THERAPIST INTERACTION BEHAVIOURS WITH CHILD OR PARENT
25. Did the THERAPIST and CHILD show a **strong rapport**? **Item Definition:** This item measures the extent to which the relationship between the therapist and the child is marked by harmony and accord (i.e., how well the child and therapist get along in the session). The rater can judge how well the dyad get along by examining the responsiveness of the child to the therapist (and vice versa) and in demonstrating mutual trust and how well they seem to enjoy each other's company in activities within intervention session.
26. Did the THERAPIST and the CHILD **work together**? **Item Definition:** This item measures the extent to which the therapist and child collaborate within the session to do activities and/or make decisions. Rather than having the child perform activities alone, the therapist collaborates with the child. The therapist could also openly suggest the child to engage or try different activities, and the child attempts to perform them. E.g., therapist participates in activities to make them a social experience rather than making the child feel as though he or she is being analyzed/tested.
27. Did the THERAPIST and PARENT show **strong rapport in their discussions**? **Item Definition:** This item measures the extent to which positive discussions are evident between the parent and therapist within an intervention session. The therapist and parent demonstrate a comfortable (or positive) dynamic and **openly** share thoughts and/or concerns about child to refine elements of the intervention in session. E.g., discussion about whether intervention goals or strategies are realistic for therapy and/or their application to the home setting.

28. Did the THERAPIST and PARENT **work together** to determine how to proceed?

Item Definition: This item measures the extent to which the therapist and parent collaborate within the session to resolve any issues that arise. E.g., discussing potential intervention strategies and/or activities for the child or family situation within the session. Interactions between parent and therapist show collaboration when discussing plans for child's care or rehabilitation.

Overall impression of observed intervention

After rating all 28 items from a videotaped intervention session, raters will be asked to give their overall impression of the intervention observed. On a scale (1 to 5), raters will report what they believe is the best overall description and quality of the observed intervention. See below.

Using the scales below, rate your judgment regarding the overall **purpose** and **quality** of the intervention provided by the therapist in the session that you have just viewed (CIRCLE ONLY ONE NUMBER). **Please base your rating taking into consideration the session as a whole.**

The <u>overall purpose</u> was to:

Change the environment or activity to enable function				Remediate impairments and improve the child's skills
1	**2**	**3**	**4**	**5**

The <u>overall quality</u> of the therapy provided within the session was:

Poor		**Average**		**Very Good**
1	**2**	**3**	**4**	**5**

Final Question:

Please indicate whether you have:

Viewed this therapist on a previous tape? _____ No _____ Yes

Appendix C

The Paediatric Rehabilitation Observational measure of Fidelity (PROF)

Video #:_____

Measure of behaviours in paediatric intervention

Video code: _____ Birth month: _ _ First three letter of mother's first name: __ __ __ Date: _____

You are part of a study that is identifying behaviours that occur within an intervention (as observed on video). Please watch the video and observe what takes place. While watching the videos, notes can be taken on one of the templates at the end of this package and/or in the right hand column of the measure.

After watching a video, answer the questions on the measure by checking the box on the scale(s) that best describes the behaviours observed in the intervention session. The measure has 28 items over four sections that focus on behaviours of (1) therapist's intervention, (2) therapist's general practice, (3) the child or parent, and (4) therapist interactions with the child/parent. Please answer all sections.

Video #:_____

Measure of behaviours in paediatric intervention

SECTION 1: THERAPIST INTERVENTION ITEMS		**Notes**
1. Did the THERAPIST focus on **remediating** the CHILD'S movements or skills? **Rate the quality of this observed behaviour**		
2. Did the THERAPIST focus on the CHILD **completing activities in a specific way** (e.g., follow a specific sequence of steps)? **Rate the quality of this observed behaviour**		
3. Did the THERAPIST focus on the **quality** of the CHILD'S movement and/or skills? **Rate the quality of this observed behaviour**		
4. Did the THERAPIST facilitate **repetition in the practice** of movements or motor skills addressed in the intervention session? **Rate the quality of this observed behaviour**		
5. Did the THERAPIST focus on **changing** the **environmental constraints** that influence the CHILD'S activity performance (e.g., environmental barriers within an activity)? **Rate the quality of this observed behaviour**		
6. Did the THERAPIST focus on **adapting an activity** to enable the CHILD to complete the activity? **Rate the quality of this observed behaviour**		

Video #:_____

Measure of behaviours in paediatric intervention

7. Did the THERAPIST provide opportunities for the CHILD to **practice activities for the purpose of completion** (i.e., not concerned with how the activity was done)? **Rate the quality of this observed behaviour**		
8. Did the THERAPIST focus on the **most efficient way that best suits the CHILD** to perform the activity? **Rate the quality of this observed behaviour**		

SECTION 2: THERAPIST GENERAL PRACTICE ITEMS **Notes**

9. Did the THERAPIST **assess** the CHILD (formally or informally)? **Rate the quality of this observed behaviour**		
10. Did the therapist **review therapy goals and/or activities** for the CHILD (and/or Parent) **to practice outside of the intervention** sessions? **Rate the quality of this observed behaviour**		
11. Did the THERAPIST **demonstrate, instruct and/or teach** the CHILD skills or strategies? **Rate the quality of this observed behaviour**		
12. Did the THERAPIST **grade** the level of **difficulty of the steps** to participate in an activity? **Rate the quality of this observed behaviour**		

Video #:_____

Measure of behaviours in paediatric intervention

		Notes
13. Did the THERAPIST offer the CHILD **activities** that were **engaging or fun**? **Rate the quality of this observed behaviour**		
14. Did the THERAPIST **praise or communicate confidence** in the CHILD'S effort or performance? **Rate the quality of this observed behaviour**		
15. Did the THERAPIST **demonstrate warmth towards** the CHILD? **Rate the quality of this observed behaviour**		
16. Did the THERAPIST alter his/her behavior (OR show flexibility) **to adapt to the needs, behaviours, or interests** of the CHILD? **Rate the quality of this observed behaviour**		

SECTION 3: CHILD/PARENT BEHAVIOUR ITEMS

17. Did the CHILD **actively participate** in the intervention session?		
18. Did the CHILD **show comfort** with the therapist?		
19. Did the CHILD **communicate his/her preferences**?		
20. Did the CHILD **show discomfort due to a personal factor** (e.g., physical pain, sleepiness, emotional lability)?		
21. Did the CHILD (or PARENT where appropriate) **practice strategies with the therapist**?		

22. Did the PARENT **participate** in the intervention session?	☐	
23. Did the PARENT **appear to be comfortable** with the therapist?	☐	
24. Did the PARENT **communicate his/her preference(s)**?	☐	

SECTION 4: THERAPIST INTERACTIONS WITH CHILD/PARENT ITEMS

25. Did the THERAPIST and CHILD show a **strong rapport**?	☐	
26. Did the THERAPIST and the CHILD **work together**?	☐	
27. Did the THERAPIST and PARENT show **strong rapport in their discussions**?	☐	
28. Did the THERAPIST and PARENT **work together** to determine how to proceed?	☐	

Using the scales below, rate your judgment regarding the overall **purpose** and **quality** of the intervention provided by the therapist in the session that you have just viewed (CIRCLE ONLY ONE NUMBER). **Please base your rating taking into consideration the session as a whole.**

The <u>overall purpose</u> was to:

Change the environment or activity to enable function				Remediate impairments and improve the child's skills
1	**2**	**3**	**4**	**5**

The <u>overall quality</u> of the therapy provided within the session was:

Poor		**Average**		**Very Good**
1	**2**	**3**	**4**	**5**

Video #:_____

Measure of behaviours in paediatric intervention

Please indicate whether you have:
Viewed this therapist on a previous tape? _____ No _____ Yes

Use the appropriate scales below to score the items on the measure **(Keep this with you while rating all items):**

Frequency Rating	N/ A No present	1 Not at all	2 A little	3 Somewhat	4 Considerably	5 Extensively
Quality Rating	N/ A Freq. rating =1	1 Poor	2 Acceptable	3 Average	4 Good	5 Very Good

Video #:_____

Measure of behaviours in paediatric intervention

WORKSHEET #1: A reference sheet to take notes while watching the videotaped intervention

1. THERAPIST INTERVENTION	2. THERAPIST GENERAL PRACTICE
3. CHILD/PARENT BEHAVIOUR	4. THERAPIST INTERACTIONS WITH CHILD/PARENT

Video #:_____

Measure of behaviours in paediatric intervention

Worksheet #2: An alternate reference sheet to take notes while watching the videotaped intervention

Intervention Behaviours	1. Focus on remediating child's skills	
	2. Child completion of activity in specific way	
	3. Focus on quality of child performance	
	4. Facilitate repetition/practice of movements or skills	
	5. Observe or change environmental constraints on child	
	6. Focus on adapting the activity	
	7. Child practice activity for purpose of completion	
	8. Focus on most efficient way to complete activity	
General Practice	9. Assess child's ability (formal/informal)	
	10. Review goals/activities practice outside intervention	
	11. Demonstrate, instruct and/or teach skills or strategies	
	12. Grade difficulty of steps to participate in an activity	
	13. Engage the child in fun activities	
	14. Praise/communicate confidence in the child's effort	
	15. Warmth towards child	
	16. Therapist flexibility or adapting behavior in session	

Video #:_____

Measure of behaviours in paediatric intervention

Client Behaviours	17. Child level of participation	
	18. Child's comfort with therapist	
	19. Child's communication of preference	
	20. Child discomfort due to personal factor	
	21. Child practices strategies with therapist	
	22. Parent participates in session	
	23. Parent's comfort with therapist	
	24. Parent communicates preference	
Intxn	25. Rapport between therapist and child	
	26. Collaboration between therapist and child	
	27. Rapport between therapist and parent	
	28. Collaboration between therapist and parent	

Appendix D:

Exploratory Factor Analysis Tables

Table D1. Exploratory Factor Analysis tables for Child Therapy Frequency versus Quality Ratings

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total
1	3.766	47.074	47.074	3.766
2	1.448	18.098	65.171	1.448
3	1.042	13.023	78.195	1.042
4	.543	6.793	84.987	
5	.510	6.380	91.367	
6	.259	3.240	94.607	
7	.230	2.870	97.478	
8	.202	2.522	100.000	

Total Variance Explained

Component	Extraction Sums of Squared Loadings		Rotation Sums of Squared Loadings		
	% of Variance	Cumulative %	Total	%of Variance	Cumulative %
1	47.074	47.074	2.666	33.325	33.325
2	18.098	65.171	2.131	26.653	59.968
3	13.023	78.195	1.458	18.227	78.195

Note. Extraction Method - Principal Component Analysis.

Rotated Component Matrix

	Component		
	1	2	3
a2	.188	.758	.228
a3	.087	.798	.241

a4	.253	.833	-.071
a5	.017	.326	.883
q2	.860	.178	.176
q3	.852	.176	.160
q4	.859	.237	-.031
q5	.597	-.029	.711

Note. Extraction Method - Principal Component Analysis;Rotation Method - Varimax with Kaiser Normalization; Rotation converged in 6 iterations.

Table D2. Exploratory factor analysis tables for context therapy frequency versus quality ratings

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total
1	3.881	48.517	48.517	3.881
2	1.378	17.222	65.739	1.378
3	1.074	13.420	79.159	1.074
4	.701	8.759	87.918	
5	.448	5.597	93.515	
6	.204	2.544	96.059	
7	.181	2.267	98.326	
8	.134	1.674	100.000	

Total Variance Explained

Component	Extraction Sums of Squared Loadings		Rotation Sums of Squared Loadings		
	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	48.517	48.517	2.890	36.125	36.125

2	17.222	65.739	1.907	23.831	59.956
3	13.420	79.159	1.536	19.203	79.159

Note. Extraction Method - Principal Component Analysis.

Rotated Component Matrix

	Component		
	1	2	3
a6	.177	.825	.033
a7	.136	.817	.324
a8	-.064	.156	.882
a9	.340	.096	.759
q6	.833	.378	-.057
q7	.661	.567	.049
q8	.848	.241	.206
q9	.933	-.053	.167

Note. Extraction Method - Principal Component Analysis; Rotation Method - Varimax with Kaiser Normalization; Rotation converged in 5 iterations.

TABLE D3. Rotated Component Matrix for General therapy behaviours across both therapies

	Component				
	1	2	3	4	5
a11	-.135	-.267	.765	.027	.366
a12	-.161	.117	.213	-.062	.823
a13	.141	.225	.606	.264	.089
a14	-.098	.067	.058	.834	.104
a15	.311	.169	.119	.732	-.143
a16	.767	-.137	.136	.375	-.022
a17	.924	.007	.048	.009	.164

a18	.257	-.150	-.020	.224	.748
q11	.082	.267	.821	-.013	.034
q12	.068	.403	.300	-.337	.657
q13	.149	.713	.435	.161	-.001
q14	.058	.659	.037	.458	.075
q15	.597	.380	.249	.482	-.100
q16	.756	.376	.151	.054	-.170
q17	.887	.210	-.153	-.078	.061
q18	.563	.638	-.013	-.067	.134

Note. Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.[a]
a. Rotation converged in 7 iterations.

Table D4. Exploratory Factor Analysis tables for frequency ratings of Child therapy versus Context therapy

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total
1	3.541	44.267	44.267	3.541
2	1.468	18.350	62.618	1.468
3	1.008	12.601	75.219	1.008
4	.485	6.057	81.276	
5	.421	5.264	86.540	
6	.384	4.800	91.340	
7	.370	4.626	95.966	
8	.323	4.034	100.000	

Total Variance Explained

Component	Extraction Sums of Squared Loadings	Rotation Sums of Squared Loadings

	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	44.267	44.267	2.644	33.052	33.052
2	18.350	62.618	1.773	22.167	55.218
3	12.601	75.219	1.600	20.001	75.219

Note. Extraction Method: Principal Component Analysis.

Rotated Component Matrix[a]

	Component		
	1	2	3
a2	.803	-.007	-.284
a3	.785	-.368	-.078
a4	.738	-.336	-.216
a5	.861	.018	.056
a6	-.235	.052	.856
a7	-.018	.298	.809
a8	-.134	.879	.078
a9	-.148	.813	.263

Note. Extraction Method - Principal Component Analysis; Rotation Method - Varimax with Kaiser Normalization; Rotation converged in 4 iterations.

Appendix E

Generalizability (G) Study Equations

Figure E1. Equations for Inter-rater reliability

$$\text{G (Inter-rater Reliability SR)} = \frac{\sigma^2_V + \sigma^2_{VI}}{\sigma^2_V + \sigma^2_R + \sigma^2_{RV} + \sigma^2_{IV} + \sigma^2_{RI} + \sigma^2_{RIV}}$$

$$\text{G (Inter-rater Reliability AVG)} = \frac{\sigma^2_V + \sigma^2_{VI}}{\sigma^2_V + (\sigma^2_R / n_R) + \sigma^2_{IV} + (\sigma^2_{RV} / n_R) + (\sigma^2_{RI} / n_R) + (\sigma^2_{RIV} / n_R)}$$

Figure E1. V = Video (facet of differentiation); I = Item (facet of generalization); R = Rater (fixed facet).

Figure E2. Equations for internal consistency

$$\text{G (Internal Consistency SR)} = \frac{\sigma^2_V + \sigma^2_R + \sigma^2_{RV}}{\sigma^2_V + \sigma^2_R + \sigma^2_{RV} + \sigma^2_{IV} + \sigma^2_{RI} + \sigma^2_{RIV}}$$

$$\text{G (Internal Consistency AVG)} = \frac{\sigma^2_V + \sigma^2_R + \sigma^2_{RV}}{\sigma^2_V + \sigma^2_R + \sigma^2_{RV} + (\sigma^2_{IV} / n_I) + (\sigma^2_{RI} / n_I) + (\sigma^2_{RIV} / n_I)}$$

Figure E2. SR = Single Rater; AVG = Averaging over a number of raters; σ^2 = Variance; n = sample.

Appendix F

Decision (D) Study Equation

The D-Study equation is set up below:

$$G \text{ (Inter-rater Reliability)} = \frac{\sigma^2_V + \sigma^2_{VI}}{\sigma^2_V + (\sigma^2_R / n_R) + \sigma^2_{IV} + (\sigma^2_{RV} / n_R) + (\sigma^2_{RI} / n_R) + (\sigma^2_{RIV} / n_R)}$$

Note: V=Video; R = Rater; I = Item; σ^2 = Variance; n = sample size

Appendix G:

Study Complete Results for G - and D - Studies

Table F1. D-Study results for G-coefficients based on number of raters by profession

Intervention type	**Occupational Therapy Raters (G-coefficient =G)**	**Physiotherapy Raters (G-coefficient =G)**
Child Therapy	G= 0.73 (5 Raters) G= 0.69 (4 Raters) G= 0.62 (3 Raters)	G= 0.75 (5 Raters) G= 0.71 (4 Raters) G = 0.65 (3 Raters)
Context Therapy	G= 0.70 (8 Raters) G= 0.66 (6 Raters)	G= 0.30 (30 Raters) G = 0.12 (10 Raters)

Table F2. D-Study Results for IRR Context Therapy Domain

Number of Raters	**G-coefficient**
1	0.33
2	0.49
3	0.6
4	0.66
5	0.71
6	0.75

Note. Items =4; across all videos n=25.

Table F3. D-Study Results for IRR Child Therapy Domain

Number of Raters	**G-coefficient**
1	0.54
2	0.7
3	0.78
4	0.82
5	0.85
6	0.87

Note. Items =4; across all videos n=25.

Table F4. D-Study Results for IRR General Attributes Domain

Number of Raters	G-coefficient
1	0.45
2	0.62
3	0.71
4	0.76
5	0.8
6	0.83

Note. Items = 20; across 19 videos.