University of Reading

Telepractice Application for the Clinical Assessment of School-Age Children Who Stutter

Thesis submitted for the degree of Doctor of Philosophy

School of Psychology and Clinical Language Sciences

Lamya Aldukair

November 2018

Table of Contents

Dedication iv

Declaration v

Acknowledgements vi

Table of Appendices viii

Table of Figures and Tables ix

List of Abbreviations xi

Abstract 1

Introduction 4

1 Overview of Stuttering 6

1.1 Introduction 6

1.2 Definition of Stuttering 6

1.3 Cause of Stuttering 8

1.4 Prevalence and Incidence 10

1.5 Sex Ratio and the Onset of and Recovery from Stuttering 12

1.6 Negative Consequences of Stuttering 12

2 Stuttering Assessment and Treatment for Children Who Stutter 15

2.1 Introduction 15

2.2 Why the Clinical Measurement of Stuttering Is Essential 15

2.3 Common Stuttering Measurements 17

2.4 Stuttering Therapies 27

3 Availability and Access to Speech and Language Therapy Services in Saudi Arabia 32

3.1 Introduction 32

3.2 Saudi Arabia 32

3.3 Availability and Access to SLT Services 33

4 Identifying the Assessment Practices of Saudi Speech and Language Therapists for School-Age Children Who Stutter 37
4.1 Introduction 37
4.2 Questionnaire development 37
4.3 Participants and procedure 38
4.4 Data analysis 38
4.5 Results 38
4.6 Discussion 45
5 SSI: Background, Procedures, and Reliability 48
5.1 Introduction 48
5.2 Background 48
5.3 Relative and Absolute Reliability of the Arabic SSI-1V: An Investigation 52
6 Telepractice 64
6.1 What is Telepractice? 64
6.2 Telepractice in Saudi Arabia 64
6.3 Internet Access 65
6.4 The History of Telepractice 66
6.5 The Scope of Telepractice 66
6.6 Telepractice in the Context of Speech and Language Therapy 67
6.7 Telepractice Delivery Models 67
6.8 Past Reviews of Telepractice Applications in the Field of Speech and Language Therapy 68
6.9 Potential Benefits and Limitations of Telepractice 71
6.10 Acceptability of and Satisfaction with the Telepractice Model 75
6.11 Telepractice Applications in Speech and Language Therapy 76
6.12 Research Problem and Aims 93

6.13 Research Questions 94
6.14 Research Hypotheses 94
7 Methodology 95
7.1 Introduction 95
7.2 Methods 95
7.3 Procedure 103
7.4 Data Analysis 108
8 Results 115
8.1 Introduction 115
8.2 Background Characteristics of the Study Participants 115
8.3 Results 117
8.4 Summary of Findings 142
9 Discussion 143
9.1 Introduction 143
9.2 Discussion of Research Hypotheses 143
9.3 Benefits, Drawbacks, and Clinical Implications of the TP Model in the Assessment of School-Age CWS 152
9.4 Conclusion 156
9.5 Considerations for SLTs Applying the TP Model for the Assessment of Stuttering 157
9.6 Limitations and Future Directions 159
References 163
Appendices 198

Dedication

This thesis is dedicated to my beloved parents,

Mohammed Aldukair and Maha Alohali.

Declaration

I confirm this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Acknowledgements

Undertaking this PhD has been a truly life-changing experience for me, and it would not have been possible without the support and guidance I received from many wonderful people.

First and foremost, I would like to express sincere gratitude to my supervisor, Dr David Ward, for his continuous support throughout my PhD, for his patience, motivation, and enthusiasm, and, most importantly, for sharing his immense knowledge of stuttering with me. His constant support and intellectual guidance, from start to finish, have been invaluable over the years, and I feel incredibly privileged to have had him as my supervisor.

My sincere thanks also go to all the children in Riyadh who participated in my project and to their families for agreeing to take part in this study.

I would like to thank the speech therapists at King Abdulaziz University Hospital – Sara Al-Hamid, Nojoud Al-Dakkan, Ayna Al-Masad, and Tahani Al-Otaibi – who gave me access to participants from their caseloads and helped with recruiting, organising appointments, and obtaining consent.

Special thanks also go to Lama Al-Ohali, who, regardless of her busy schedule, devoted her time to assess children and score samples for the main study, not to mention her support, friendship, and great heart. I would also like to thank Sara Al-Hamid, Atheer Al-Humoud, and Ghada Al-Dukair for scoring samples for the SSI-IV reliability study.

I would like to thank all the friends I have met in London throughout these years, who have provided me with the necessary distractions from my project and made my stay in London unforgettable.

Special thanks to my friend Madhawi Al-Taib, with whom I shared this PhD journey; thank you for your generous help and support and for the many lively discussions and good times we have had.

I have been blessed with a very loving and supportive family. I would like to thank my parents, Mohammed Aldukair and Maha Alohali, who were my true inspiration and were the backbone of my PhD. Without their encouraging words and support, it would not have been possible to obtain this degree. They are the most important people in my world, and I dedicate this thesis to them.

I would also like to thank my three amazing sisters – Sarah, Dana, and Ghada – to whom I find it difficult to express my gratitude because it is so boundless. They are my most enthusiastic cheerleaders, and they always believed in me and encouraged me to follow my dreams.

Last but not least, I would like to thank my husband, Abdulaziz Al-Saleh, whom I met during the final year of my studies; thank you for the final push and unconditional love and support you gave me. I love you, and I cannot wait to spend the rest of my life with you, PhD-free!

Table of Appendices

Appendix 1 Questionnaire to Explore the Assessment Practice of Saudi Speech and Language Therapists 199

Appendix 2 Procedure for Applying and Scoring the SSI-IV 206

Appendix 3 The Arabic Version of the SSI-IV 209

Appendix 4 Information Sheet 210

Appendix 5 Institutional Review Board (IRB) Approval 213

Appendix 6 Parents' Consent Form 214

Appendix 7 Child Consent Form 216

Appendix 8 Questionnaires 216

Appendix 9 Types of ICC Models 225

Appendix 10 Summary of Audio and Visual Ratings of Sessions 226

Appendix 11 Additional Parents' Comments 229

Appendix 12 Children's Additional Comments 231

Table of Figures and Tables

Figure 1 Map of Saudi Arabia 33
Figure 2 Screenshot of Reading Passage Presented and Shared Online via WebEx Whiteboard from the Participant's End 100
Figure 3 Picture of a TP-Led Assessment from the Leading SLT Site 105
Figure 4 Structure of a Bland-Altman Plot with Explanation of Elements 112
Figure 5 Box-Plots of the Scores of %SS in Face-to-Face and TP Environments before Conducting Sensitivity Analysis..... 119
Figure 6 Box-Plots of the Scores of %SS in Face-to-Face and TP Environments after Conducting Sensitivity Analysis..... 119
Figure 7 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of %SS before Conducting Sensitivity Analysis..... 120
Figure 8 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of %SS after Conducting Sensitivity Analysis..... 120
Figure 9 Box-Plots of the Scores of %SS Subcomponent in Face-to-Face and TP Environments. 121
Figure 10 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of the %SS Sub-Score. 122
Figure 11 Box-Plots of the Scores of the Duration Subcomponent in Face-to-Face and TP Environments. 123
Figure 12 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of the Duration Subcomponent. 124
Figure 13 Box-Plots of the Scores of Physical Concomitants in Face-to-Face and TP Environments. 125
Figure 14 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of Physical Concomitants..... 126
Figure 15 Box-Plots of the SSI-IV Total Scores in Face-to-Face and TP Environments..... 127
Figure 16 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of Total Scores. 128
Figure 17 Box-Plots of SR in Face-to-Face and TP Environments 129
Figure 18 Bland–Altman Plot of the Difference in TP and Face-to-Face SR Scores against the Mean Scores of SR..... 130

Figure 19 An Example of Picture Plates Used to Elicit Spontaneous Speech Samples 206

Table 1 Summary of Respondents' Demographic Characteristics 39
Table 2 Assessment Components Included by Saudi SLT in a Stuttering Assessment..... 43
Table 3 Inter-Rater Relative Reliability of the SSI-IV between the Four Judges 56
Table 4 Inter-Rater Absolute Reliability of the SSI-IV between the Four Judges 56
Table 5 Intra-Rater Relative Reliability of the SSI-IV between the Four Judges 57
Table 6 Intra-Rater Absolute Reliability of the SSI-IV of Judge 1 58
Table 7 Intra-Rater Absolute Reliability of the SSI-IV of Judge 2 59
Table 8 Intra-Rater Absolute Reliability of the SSI-IV of Judge 3 59
Table 9 Intra-Rater Absolute Reliability of the SSI-IV of Judge 4 60
Table 10 Equipment Description and Use 102
Table 11 Background Characteristics of the Study Participants..... 116
Table 12 LOAs of %SS and SR..... 133
Table 13 LOAs of SSI-IV Subcomponents 133
Table 14 Inter-/Intra-Rater Reliability of Face-to-Face Ratings of the %SS and SR: ICC (95% CI). 134
Table 15 Inter-/Intra-Rater Reliability of TP Ratings of the %SS and SR: ICC (95% CI) ... 134
Table 16 Inter-/Intra Rater Reliability of TP Ratings of the SSI-IV Subcomponents: ICC (95% CI) 134
Table 17 Results of Pre- and Post-Assessment Questionnaires of Parents..... 140
Table 18 Results of Pre- and Post-Assessment Questionnaires of Children 141

List of Abbreviations

ASHA	American Speech Language Hearing Association
BAB	Behaviour Assessment Battery
BCL	Behavioural Checklist
BDAE	Boston Diagnostic Aphasia Examination
BNT	Boston Naming Test
CAT	Communication Attitude Test
CELF	Clinical Evaluation of Language Fundamentals
CP	Camperdown Program
CWS	Children who Stutter
DPOAE	Distortion Product Otoacoustic Emissions
DV	Dependent Variable
EOWPVT	Expressive One-Word Picture Vocabulary Test
f2f	face to face
GFTA	Goldman-Fristoe Test of Articulation
HIPAA	Health Insurance Portability and Accountability Act
ICC	Intra-class Correlation
ICD	International Classification of Disease
ICF	International Classification of Functioning, Disability and Health
IV	Independent Variable
KAUH	King Abdulaziz University Hospital
Kw	Quadric Weighted Kappa
LOA	Limits of Agreement
LP	Lidcombe Program
MOH	Ministry of Health

OAESES	Overall Assessment of the Speaker's Experience of Stuttering
PCIT	Parent-Child Interaction Therapy
PLS	Preschool Language Scale
PPVT	Peabody Picture Vocabulary Test
PWS	People who Stutter
QoL	Quality of Life
REEL	Receptive-Expressive Emergent Language Test
SKOLD	Screening Kit of Language and Development
SLT	Speech and Language Therapist/Therapy
SPM	Syllables per Minute
SR	Severity Rating
%SS	Percentage of Syllables Stuttered
SSI	Stuttering Severity Instrument
TP	Telepractice
WASSP	Wright and Ayre Stuttering Self-Rating Profile
WHO	World Health Organization
WPM	Words per Minute

Abstract

Introduction: In Saudi Arabia, despite the country's relatively high expenditure on healthcare, where investment in health services by the Ministry of Health (MOH) has witnessed a rise from 6.49% of total government expenditure in 2010 to 7.27% in 2015 (Ministry of Health, 2015), access to speech and language therapy services is limited. Addressing the public's inequitable access to resources is considered a major issue in Saudi Arabia's healthcare system (Alkabba et al., 2012), where regular healthcare provisions, including speech and language therapy services, are not available to many of the country's children living either in major cities or in rural or remote areas.

One way to improve access for those who may otherwise be deprived of speech and language therapy assessment or treatment is through telepractice (TP). Studies investigating the use of TP services with children and adults who stutter have shown similar therapeutic outcomes when compared to traditional face-to face delivery (e.g. Carey et al., 2010; Lewis et al., 2008; O'Brian et al., 2014; Wilson et al., 2004); however, there is currently no evidence as to the viability of TP stuttering assessments for any age group (Lowe et al., 2013). The present research therefore investigates the validity and reliability of using a TP application to assess stuttering behaviour in school-age children.

Methodology: Prior to commencing the main study, a questionnaire study was conducted to understand the assessment practices followed by speech language therapists (SLTs) in Saudi Arabia, particularly their assessment practices in the field of stuttering, so as to aid in the selection of the most relevant assessment activities related to stuttering in Saudi clinical settings. The Stuttering Severity Instrument (Fourth Edition) (SSI-IV) (Riley, 2009), Percentage of Syllables Stuttered (%SS), and the Severity Rating scale (SR) were chosen as assessment measures for the main study. Further, a study to investigate the reliability of the SSI-IV was conducted. The text of the SSI-IV examiner record form was translated into Arabic according to the Process of Translation and Adaptation of Instruments guidelines (WHO, 2015), and its relative and absolute reliability was investigated. The main study consisted of thirty children who stutter (CWS), aged between 6 and 15 and recruited from a speech therapy clinic in King Abdulaziz University Hospital in Riyadh. Children were divided into two testing conditions, either face-to-face-led or TP-led testing, and were administered the SSI-IV, %SS, and SR either face-to-face or via TP using two laptops with

webcams, video conference software, and a broadband internet connection. Agreement and reliability of scoring in the two testing conditions were reported, in addition to questionnaires exploring children's and their caregivers' perceptions both before and immediately after the TP assessment being carried out.

Results: Findings of the questionnaire study revealed that most respondents assess various dimensions of stuttering (covert and overt). With regard to the assessment of overt features of stuttering, most respondents reported the use of formal and quantifiable measures, such as the %SS and SSI. Respondents reported using a non-formal approach for the assessment of attitudes and quality of life through interviews. As for the reliability of the translated SSI-IV, high to very high relative intra- and inter-rater reliability was reported. Regarding absolute reliability, intra-rater reliability was more favourable when compared to inter-rater reliability. Finally, for the main study, 30 sessions were successfully carried out, with results revealing unsatisfactory levels of agreement when applying the Bland and Altman (1987) method. However, when discrepancies were found, it was comparable to those from traditional face-to-face studies. Generally, parents and children exhibited high levels of satisfaction and held a positive view regarding TP pre- and post-assessment.

Discussion and Conclusion: Results of the questionnaire distributed to Saudi SLTs to explore their views regarding stuttering assessment highlight the need for SLTs to receive more training in the domain of stuttering in order to make them aware of up-to-date assessment tools and teach them to adapt well-established Western tools for the Arabic language or develop authentic assessment tools and scales for the target population. With regard to the SSI-IV reliability investigation, the results suggest that the Arabic version of the SSI-IV can reliably be used by SLTs for research and clinical purposes. However, the compromised absolute inter-rater reliability suggests that SSI-IV may not be appropriate for use when identifying minor changes in stuttering severity within an individual across time, when different SLTs measure the stuttering severity of the same client. The results of the main study suggest that conducting a stuttering assessment via TP for school-age CWS is feasible, reliable, and valid. Although the TP model posed some challenges for executing the assessment, mainly technological challenges, high levels of satisfaction were documented in the parents and children. The TP system and equipment chosen in this study has provided a basis for the delivery of TP stuttering assessment in a clinical setting, thus addressing the barriers to access that are present for school-age CWS in Saudi Arabia. These results are

preliminary, but they can be seen as a building block for future research in TP assessment studies for CWS.

Introduction

This PhD research study is primarily concerned with the assessment of stuttering through an alternative service delivery model called telepractice (TP). Thus, we will concentrate, in depth, on a range of theoretical and clinical issues related to stuttering as well as telepractice.

The inspiration for conducting this study stemmed from the continuing challenges speech and language therapists in Saudi Arabia encounter as a result of the barriers to access and service limitations preventing children from receiving assistance from specialised speech and language therapy services. I worked as a speech and language therapist for a year prior to commencing my PhD, and I often had children who stuttered in my caseload; I have seen participants travel what is equivalent to a ten-hour drive just to receive a speech and language therapy consultation or follow-up sessions. Even when children and their parents are highly motivated to receive therapy, the availability of regular sessions may be an obstacle, thus compromising the speech and language therapy services given, which, for school-age children, is exacerbated by the fact that speech and language therapy services are completely lacking in school settings.

As an advocate of evidenced-based practice and the belief that all levels of speech and language therapy, clinical practice, and service delivery should be informed by scientifically-derived findings (Reilly, 2004), I felt that it was important to convert this clinical need into a question that could be answered through research, which led me to developing an interest in exploring alternative service delivery models to help expand speech and language therapy services throughout the country.

This thesis is comprised nine chapters, the first six of which are literature review chapters, including two small studies, which precede the main study. In Chapter One, an overview of stuttering is provided, encompassing issues related to definitions, causes, incidence, prevalence, sex ratio, onset, recovery, and finally the consequences of stuttering throughout an individual's life. In Chapter Two, the focus is on discussing various stuttering measurements, along with the therapeutic options for children who stutter. Since this thesis is concerned with stuttering assessment, particularly the assessment of stuttering via TP, the most common stuttering measurements will be critically reviewed. Chapter Three sheds light on the availability of and access to SLT services in Saudi Arabia. In this chapter, Saudi Arabia's demographics are outlined, and then the availability of and access to SLT services

are discussed in order to pinpoint how difficult it is for some children and their families to access SLT services in a timely and effective manner. Chapter Four describes a study conducted to understand the assessment practices followed by SLTs in Saudi Arabia, particularly their assessment practices in the field of stuttering, so as to aid in the selection of the most relevant assessment activities related to stuttering in Saudi clinical settings for the main study. Following this, Chapter Five is dedicated to describing the background and functioning of the SSI-IV, together with a discussion of its relative strengths and weaknesses, before concluding with a study that outlines the processes and procedures involved in its translation into Arabic, together with a small pilot study that tests its relative and absolute reliability. In Chapter Six, telepractice, a potential means for improving access to SLT services, is reviewed, in addition to presenting the research question, aims, and hypotheses for the main study. In Chapter Seven, the study's methodology is described. The findings of the study are presented in Chapter Eight, whereas Chapter Nine considers the findings of the main study in light of the existing literature. This final chapter also discusses the benefits, drawbacks, and clinical implications of the TP model in the assessment of school-age children who stutter, in addition to outlining the considerations for speech and language therapists who wish to apply the TP model to the assessment of stuttering, the conclusions of this study, its limitations, and future research directions.

1 Overview of Stuttering

1.1 Introduction

Stuttering refers to a communication disorder characterised by interruptions to the flow of speech that occur in the form of either repetitions, prolongations, abnormal stoppages, or blocks of sounds and syllables (Craig et al., 2002). Depending on their aetiology, there are different types of stuttering, such as neurogenic, psychogenic, or developmental. Our focus in this thesis is on developmental stuttering; thus, when we refer to stuttering, it is the developmental type, which commonly arises during childhood, particularly in the preschool years (Guitar, 2006). In this chapter, an overview of stuttering will be provided, including its definition, causes, incidence, prevalence, sex ratio, onset, and recovery, in addition to the consequences of stuttering throughout an individual's lifespan.

1.2 Definition of Stuttering

There is no single definition of stuttering, as many attempts have been made to define the disorder but with little consensus. In other words, there is no definitive method for determining whether a child stutters or not (Howell, 2011). If a child is stuttering, there are signs of this condition in the child's speech attempts (Howell, 2011). In addition, children who stutter (CWS) exhibit social and emotional struggles (Langevin, Packman & Onslow, 2010). Such indicators, in a child's speech and other features, that may help identify whether the child is stuttering or not can be found in some of the early definitions of stuttering.

Van Riper (1982) classified stuttering as a motor speech disorder: 'Stuttering occurs when the forward flow of speech is interrupted by a motorically disrupted sound, syllable or word or by the speaker's reaction thereto' (p.52). According to Bloodstein (1995), stuttering is the breaking up of the fluency or rhythm of speech with blanks or specific interruptions, whereas Wingate (1964) argued that stuttering occurs as a consequence of the peripheral speech mechanism being uncoordinated. Wingate's (1964) definition is regarded by experts in the field as the most detailed and systematic definition of stuttering, given that it consists of seven elements categorised under three headings as outlined below (p.488):

I. (a) Disruption in the fluency of verbal expression, which is (b) characterized by involuntary, audible or silent, repetitions or prolongations in the utterance of short speech elements: namely, sounds, syllables, and words of one syllable.

These disruptions (c) usually occur frequently or are marked in character and (d) are not readily controllable.

II. Sometimes the disruptions are (e) accompanied by accessory activities involving the speech apparatus, related or unrelated body structures, or stereotyped speech utterances. These activities give the appearance of being speech-related struggle.

III. Also, there are not infrequently (f) indications or report of the presence of an emotional state, ranging from a general condition of 'excitement' or 'tension' to more specific emotions of a negative nature, such as fear, embarrassment, irritation, or the like. (g) The immediate source of stuttering is some incoordination expressed in the peripheral speech mechanism; the ultimate cause is presently unknown and may be complex and compound.

The International Classification of Diseases (ICD) defines stuttering as

> 'speech that is characterized by frequent repetition or prolongation of sounds or syllables or words, or by frequent hesitations or pauses that disrupt the rhythmic flow of speech. There may be associated movements of the face and/or other parts of the body that coincide in the time with the repetitions, prolongations, or pauses in speech flow.' (World Health Organization, 1993: p.227).

Similarly, in the DSM-V (2013), stuttering criteria are defined as the frequent occurrence of one or more of the symptoms below, accompanied by problems in the fluency and timing of the speech sound and syllable repetitions:

1- sound prolongations;

2- broken words (e.g. pauses within a word);

3- audible or silent blocking (filled or unfilled pauses in speech);

4- circumlocutions (word substitutions to avoid problematic words);

5- words produced with an excess of physical tension; and

6- monosyllabic whole-word repetitions (e.g. 'I-I-I I see him').

1.3 Cause of Stuttering

The cause of stuttering is still inconclusive (Packman, Code & Onslow, 2007). However, there is compelling evidence that stuttering encompasses genetic and neurophysiological factors that play a role in the emergence of stuttering (Packman, 2012).

There is strong evidence for the role of genetics in stuttering (for a review, see Kraft & Yairi, 2012). In family incidence studies, which are among the first to study the possibility of a genetic contribution in stuttering and consist of basically counting the percentage of relatives of a stuttering client who stutter, Yairi and Ambrose (1999) revealed that 20–47% of people who stutter (PWS) report at least one family member who stutters. However, this method has been criticised as being inaccurate and misleading due to the fact that family size or/and familial class were not considered in the studies (Yairi & Ambrose, 2013). As a more refined research study, twin studies suggest a high concordance of stuttering for monozygotic twins of up to a 50 or 60% rate of incidence (Felsenfeld et al., 2000; Howie, 1981). This is an indication that genes do not *cause* stuttering, as twins may both begin stuttering and one may recover while the other does not (Smith & Weber, 2016). In this sense, infants are not born programmed to stutter; rather, stuttering behaviours emerge during development, highlighting the role of extrinsic factors (Smith & Weber, 2016). In family aggregation studies, one or more possible genes responsible for stuttering have been identified (Andrews et al., 1991; Kidd et al., 1981). Recent studies have revealed contributing genes to stuttering, as different genes on various chromosomes have been linked to stuttering (Kang et al., 2010; Lan et al., 2009; Riaz et al., 2005; Shugart et al., 2004; Suresh et al., 2006; Wittke-Thompson et al., 2007).

Initially, neurophysiological studies were only focused on the study of adults. These studies found that structural and functional differences exist between the brains of PWS and fluent controls. Structural differences were found in grey and white matter and in left–right asymmetries (Beal et al., 2007; Choo et al., 2011; Foundas et al., 2001; Jäncke, Hänggi & Steinmetz, 2004; Kell et al., 2009; Lu et al., 2010; Somer et al., 2002; Song et al., 2007; Watkins et al., 2008). Moreover, functional differences include increased activation in the right hemisphere speech regions (Braun et al., 1997; De Nil & Kroll, 2001; Fox et al., 2000; Salmelin et al., 1998, 2000). These structural and functional differences comprise both motor and auditory brain regions (Watkins et al., 2008). Also, following intensive speech treatment,

the neural activation patterns in stuttering adults showed a reduction in over-activation and a shift towards a more left-lateralised activation pattern (De Nil et al., 2003).

As mentioned earlier, neurophysiological studies of stuttering initially focused on adults. However, studies have since emerged seeking to understand the neurophysiological nature of children who stutter. Current brain research involving CWS from the age of six onwards has shown differences from earlier studies addressing adults, where CWS had less grey matter volume in the bilateral inferior frontal gyri and middle temporal gyrus when compared to fluently speaking children (Beal et al., 2013; Chang et al., 2008). More recently, Chang et al. (2015), employing a larger sample that included considerably younger children (from three years old upwards), also found that the level of fractional anisotropy in tracts interconnecting auditory-motor areas and tracts that support skilled movement control was differentiated in CWS when compared to children who did not. Those with low fractional anisotropy had more severe stuttering than those with high fractional anisotropy. Additionally, there were statistically significant differences in sex among CWS in the patterns of white matter development; in this sense, studies have found differences between females and males in terms of brain structure (Shaywitz et al., 1995) and between those with persistent stuttering and those who recovered (Chang et al., 2008).

Many causal theories have been proposed over the years (for a review, see Bloodstein & Ratner, 2008; Packman & Attanasio, 2017; Yairi & Seery, 2015), and it has been suggested that causal theories of stuttering can be divided into two categories; the first are theories that offer an explanation of the underlying causes of stuttering, i.e. distal causes, whereas the second are theories that offer an explanation of the individual moments of stuttering, i.e. proximal causes (although some theories address both categories) (Bloodstein & Ratner, 2008; Packman & Attanasio, 2017). There is general agreement among researchers that stuttering is multifactorial and encompasses a genetic predisposition with a combination of environmental, neurological, linguistic, motor, and learned behavioural aspects (Packman, 2012) and that individuals who stutter have different combinations of these factors, due to the unique nature of symptoms (Millard, Nicholas & Cook, 2008). Arguably, the two most influential multifactorial models are the Demand and Capacities model (Starkweather, 1987) and the Dynamic Multifactorial model (Smith & Kelly, 1997). The Demands and Capacities model assumes that stuttering is promoted when demands on fluency exceed the child's capacities for fluency, with demands and capacities being categorised as motoric, emotional,

linguistic, or cognitive (Starkweather, 1987), whereas the Dynamic Multifactorial model states that 'stuttering emerges from the complex, nonlinear interaction of many factors. No single factor can be identified as "the cause" of stuttering' (Smith & Kelly, 1997: p.209).

Recently, Packman and Attanasio proposed a three-factor model of stuttering, referred to as the 'P&A Model' (Packman, 2012). While the aforementioned models adopt a multi-factorial concept to explain stuttering, the P&A model goes one step further and explains these factors operationally (Packman, 2012). In this regard, it attempts to differentiate between the underlying cause of stuttering (i.e. the distal cause) and the cause of individual moments of stuttering (i.e. the proximal cause). These three factors are:

> '(1) a deficit in neural processing underpinning spoken language, which renders the speech production system unstable and prone to perturbation, (2) triggers, which are some inherent features of spoken language that increase the motoric task demands on that system, and (3) modulating factors, which determine the trigger.' (Packman, 2012: p.227)

1.4 Prevalence and Incidence

Epidemiological studies (for example, prevalence and incidence studies) of communication disorders are important as they aid in identifying the proportion of the population affected and thus in planning appropriate services for the affected population (Bryson, 1996). Incidence refers to the number of new cases of a certain disorder that occurs at a given time, whereas prevalence refers to the number of cases in a population at a given time (Onslow, 2017). In the case of stuttering, it is well known that the lives of PWS are affected by several contexts, which can relate to education, employment, or psychological wellbeing (see Section 1.6 for more details). This is where the importance of epidemiological study arises as it aids in an estimation of the proportion of the population who may be at risk of these kinds of consequences and in estimating the resources needed to support PWS through these challenges throughout their lifespans (Abou Ella et al., 2015).

Stuttering occurs in every language and culture studied (Bloodstein, 1995). It has been generally agreed that the prevalence of stuttering is around 1% (Bloodstein & Ratner, 2008), which suggests that around 70 million people around the world suffer from stuttering. However, studies have reported varying estimates of prevalence, likely attributable to the fact that estimates are dependent on a child's age (Yairi & Ambrose, 2013). In this sense, the

prevalence of stuttering tends to be higher in children under the age of six when compared to older children (Yairi & Ambrose, 2013). Prevalence for children under five years old has been estimated in studies as ranging between 2.2 and 5.6% (McLeod & Harrison, 2009; Okalidou & Kampanaros, 2001; Proctor et al., 2008). In contrast, the prevalence rates in studies that included samples of older children ranged from .33 to 1.60 (Boyle et al., 2011; McKinnon, McLeod & Reilly, 2007; Van Borsel et al., 2006). Lifetime incidence has been suggested as being 5% (Yairi & Ambrose , 2005), although this figure has been regarded by some as an underestimation. In this regard, according to Bloodstein and Ratner (2008), 'it would seem that a plausible figure for the lifetime incidence of all those who at some time either consider themselves or are considered by their parents to be stutterers is at least as high as 10 percent' (p.91). In addition, Yairi and Ambrose (2013) concluded that the lifetime incidence of stuttering may be higher, and the reason for the lower lifetime incidence rate may be due to underdiagnoses, self-reporting, and unsatisfactory subject selection. Reilly et al. (2009) reported that, from a cohort of 1,910 children, the 36-month cumulative stuttering incidence was 8.5% by the age of three, whereas, in a later study, this figure increased to 11.2% by the age of four (Reilly et al., 2013). These estimates of incidence and prevalence highlight the importance of providing sufficient SLT services for the stuttering population, given how common it is. Other studies of early stuttering have reported incidence rates ranging from 3 to 17% (Dworzynski et al., 2007; Felsenfeld et al., 2000; Månsson, 2000). However, this is unsurprising, as the incidence data is not clear-cut due to, for example, many studies using different definitions of stuttering and different studies following different methods for calculating incidence (e.g. retrospective and prospective recall data collection and the identification of stuttering using clinical evaluation and parent report) (Guitar, 2006; Kefalianos et al., 2017). In the Middle East, very little data is available concerning the epidemiology of stuttering. One study, carried out in 1974 in Egypt, involved 8,459 children aged between 6 and 12 from governmental and private schools in Eastern Cairo; school records and personal interviews were the tools used to determine stuttering, and the prevalence rate in this study was reported as 0.93% (Okasha et al., 1974). In a more recent study, it was found that the prevalence of stuttering among 8,765 primary school children in Cairo was 1.03% (Abou Ella et al., 2015). It should be noted that epidemiological studies of stuttering in Saudi Arabia or even the Gulf area are non-existent.

1.5 Sex Ratio and the Onset of and Recovery from Stuttering

Studies indicate that stuttering affects males more than females, especially in the case of adults (Yairi & Ambrose, 2013). With regard to the male/female ratio, it differs across one's lifespan; in young children, a ratio of 2:1 has been reported, which increases to up to 5:1 in adults (Craig et al., 2002; Yairi & Ambrose, 1999).

Regarding onset, studies indicate that stuttering usually begins in early childhood and that onset occurs most often between the ages of two and five, with 90% of cases of stuttering starting before five years of age (Howell, Davis & Williams, 2008; Månsson, 2000; Yairi & Ambrose, 2005). Bloodstein and Ratner (2008) conducted a review of stuttering onset gathered via parental recall. The study included ten reports, with a mean of 3.3 years, and four reports of a median onset of 3.9 years, suggesting that stuttering commonly occurs during some point around the age of four. If stuttering is diagnosed during childhood, the chance of spontaneous recovery is regarded as high, hence reports of longitudinal studies showing recovery rates of 65% in the age range of two to five years and ten months (Ryan, 2001), and 74% in the age range of two to four years and eleven months (Yairi & Ambrose, 1999). It has been noted that recovering from stuttering after reaching teenage years is rare (Ward, 2018). Such figures, therefore, imply that at least 20% of children who stutter will persist in doing so (Cavenagh et al., 2015).

1.6 Negative Consequences of Stuttering

Research studies indicate that CWS are aware of their stuttering as early as their preschool years (Langevin, Packman & Onslow, 2010), and it has been reported that nearly half of preschool CWS are aware of their stuttering (Boey et al., 2009). It was found that children as young as three years old demonstrate behaviour such as crying, laughing, blushing, or making comments such as, 'I can't talk' (Bloodstein & Ratner, 2008) and that such awareness can lead to negative reactions in children towards their stuttering. Studies have compared communication attitudes between preschool CWS and non-stuttering children, with results revealing that preschool-age CWS have more negative attitudes towards communication when compared to their non-stuttering peers (Clark et al., 2012; Vanryckeghem & Brutten, 2007; Vanryckeghem et al., 2005). Understandings of the awareness of preschool children who stutter and their negative communication attitudes have led to the development of standardised attitudinal scales that are commonly used with older children and adults, one example of which is the KiddyCAT scale (Vanryckeghem & Brutten, 2007) (see Chapter 2,

Section 2.3.2.1, for more details), which is used to investigate the attitudes towards communication of preschool children aged between three and six. A study by Cardell (2010) revealed that 63 CWS had more negative communication attitudes when compared to 43 matched children who did not stutter.

Langevin, Packman and Onslow (2010) investigated the impact of stuttering on preschool children and their parents, with a sample size including 77 parents who reported that their children exhibited frustration (81.8%), low self-esteem (32.2%), and compromised mood (42.9%). It was also reported that children tended to communicate less after the onset of stuttering (25%), whereas 23.4% of parents reported that their children were reluctant to communicate. In addition to the negative impact on children, parents reported social consequences of stuttering, including their children being subject to bullying by their peers. They also reported that, when their children interacted with peers, they were not given enough time to finish their sentences and were subject to interruptions during conversations; further, when a child was having a stuttering moment, their peers withdrew.

When children reach their school years, in general, they are aware of their stuttering (Boey et al., 2009), understand the stuttering features that encompass their speech, and can describe how their stuttering influences their communication abilities and participation in everyday life (Millard & Davis, 2016). In addition, with an increase in awareness, they consequently become more mindful that their speech is distinctive from their peers (Cook & Botteril, 2005). Further, at this stage, evidence of bullying increases, and, to cope with this, avoidance strategies to conceal or minimise their stuttering emerge (Cook & Botteril, 2005). In the field of stuttering, the term ‘avoidance’ can be defined as behaviours that are learned from disfluency anticipation and the negative experiences associated with stuttering that facilitate escape from such moments or experiences (Guitar, 2006). These strategies are regarded as coping strategies or as a method of protection from the emotional harm of stuttering (Cream et al., 2003). The use of these avoidance behaviours may arise during specific sounds, words, topics, or speaking situations (Cream et al., 2003; Guitar, 2006). The impact of stuttering from a social perspective and the behaviour of school-age children was explored by Davis, Howell and Cooke (2002), who reached the conclusion that children who stutter are more often rejected by peers, more subjected to bullying, less likely to attend social gatherings, and less likely to be chosen as leaders. This implies that stuttering has broader consequences beyond impaired functioning, whereby it may influence a child’s self-esteem, confidence, and

affective state (Langevin, Packman & Onslow, 2010). For adolescents who stutter, the impact of stuttering is further exacerbated by the transition through a very confusing developmental period, both biologically and psychosocially. Changes through this phase encompass new roles, new social groups, new privileges, and a desire to explore one's developing life. However, the compromised ability to communicate to an optimum level as a consequence of stuttering or the associated fear may have an impact on this phase, resulting in a negative effect that may continue beyond this age group (Erickson & Block, 2013; McAllister, Collier & Shepstone, 2013). Moreover, adolescents who stutter may perceive themselves to be poor communicators, consequently affecting their self-esteem (Blood & Blood, 2004).

In the case of adults, stuttering is often linked to social maladjustment, under-achievement, and problems with verbal communication (Bloodstein & Ratner, 2008; Craig & Calver, 1991; Crichton-Smith, 2002; Hayhow, Cray & Enderby, 2002). When comparing adults who stutter with adults who do not, there is a higher prevalence of anxiety-related mental health disorders (Craig & Tran, 2005; Ezrati-Vinacour & Levin, 2004; Iverach et al., 2009). It was reported that almost 50% of adults who stutter qualify for the diagnosis of social phobia (Iverach et al., 2009; Kraaimaat, Vanryckeghem & Van Dam-Baggen, 2002; Menzies et al., 2008). It is thought that social phobia in adults who stutter may be a consequence of negative social encounters during school years (O'Brian et al., 2011). Adults who stutter may encounter problems in educational attainment and in vocational settings, as a negative relationship between educational attainment and stuttering severity has been reported (O'Brian et al., 2011). In addition, people who stutter may be hindered in employment settings because of how their supervisors or colleagues view them in terms of job performance, promotion opportunities, competency, or intelligence (Klein & Hood, 2004; Palasik et al., 2012). Further, seeking therapy may create a financial burden on people who stutter as, due to the complex nature of stuttering, effective therapy requires a relatively long-term commitment (Zebrowski, 2016). In an Australian study by Blumgart, Tran and Craig (2010), it was estimated that, on average, treatment for stuttering could cost $5,500 AUD for an adult who stutters in a five-year period.

2 Stuttering Assessment and Treatment for Children Who Stutter

2.1 Introduction

While the importance of the clinical measurement of stuttering is generally agreed upon, it is possibly one of the most controversial subjects in the field of dysfluency. This difficulty may be related to the variability of the disorder since, as described in Chapter 1, there is no unified definition of stuttering. Usually, stuttering is identified by the assessment of two basic features of the disorder: 1) the assessment of motor speech activity by assessing fluency counts and speech rates, usually referred to as the *overt* features of the disorder and 2) assessment of cognitive aspects represented in attitudinal questionnaires, usually referred to as the *covert* features. It can be said that the former assessment method deals with the stuttering itself, whilst the latter relates to the person who stutters (Ward, 2018). The focus on the assessment of overt features has been extensively criticised and debated, based on the notion that stuttering is much more than stuttering counts, in addition to the reliability and validity of such assessments having been questioned. On the other hand, the possibility of assessing covert features of stuttering objectively (i.e. using scales and formal questionnaires) has also been challenged. A question that therefore needs to be asked is, given all the controversy around the measurement of stuttering and given that is a difficult task, why is it essential to assess stuttering? However, regardless of the drawbacks of measurements of stuttering, it is crucial to apply them in order to gain a holistic perspective of the disorder and its impact on an individual. This chapter will begin by highlighting why the clinical measurement of stuttering is important, before reviewing the most common measurements of stuttering and discussing common therapy options for children who stutter.

2.2 Why the Clinical Measurement of Stuttering Is Essential

According to Onslow (2017), clinical measurement is essential for the following reasons:

- Assessment

Clinical measurement during a client's first visit to a clinic is important in order to provide a formal way of reporting the nature of the stuttering and the severity with which it is impacting the individual. The impact may be behavioural, which is associated with stuttering moments

and how frequently they occur, or non-behavioural, which is specifically associated with the assessment of anxiety.

- Communication with clients

It is thought that clinical measurement may facilitate communication between the speech and language therapist (SLT) and the clients and/or parents by offering a common language to discuss the client's stuttering severity. For example, when the client reports that he/she had a measurement score of five on scale of 1–9 on a particular day, the SLT will have an idea of the client's stuttering severity during that identified day. This is considered important as it helps to evaluate whether the planned treatment was effective or not.

- Stating treatment goals

In any treatment programme in the field of speech and language therapy, goal setting is the first building block toward successful treatment. Clinical measurements offer a formal description of those goals and therefore provide accountability for treatments, which may be essential if the treatments are funded (e.g. by insurance companies).

- Assessing progress toward treatment goals

Clinical measurement is not confined to the first visit only. Rather, it is a dynamic process and an essential part of treatment as it offers evidence of progress. In addition, in the case where there is no progress, measurements may provide an indication of what aspects are affecting the success of therapy and are therefore in need of modification.

- Managing maintenance of treatment gains

Stuttering can be managed successfully within the clinical environment. However, the difficulty of maintaining these positive changes and generalising them throughout everyday life may pose a challenge. Most stuttering therapy programmes have a maintenance period in order to ensure that a relapse does not occur and that positive gains are maintained. This is where clinical measures are useful, as they allow the monitoring of post-treatment progress and help in the detection of relapse.

- Keeping track of stuttering severity change

It is well documented that stuttering is variable. The frequency of a speaker's disfluencies, as well as their intensity and duration, vary markedly from situation to situation and from day to

day (Constantino et al., 2016; Karimi et al., 2013). It is therefore important to keep track of such changes through applying clinical measurements.

2.3 Common Stuttering Measurements

2.3.1 Assessment of overt stuttering behaviours

2.3.1.1 Percentage of Spoken Syllables that are Stuttered (%SS)

Recording percentage of syllables stuttered (%SS) may be regarded as the most commonly used method used to gather data objectively, and by some as the gold standard for stuttering assessment (O'Brian et al., 2015). It is usually referred to as the stuttering frequency, relating to the percentage of spoken syllables that are stuttered, commonly abbreviated as %SS and widely incorporated in clinical contexts and research trials (Cordes & Ingham, 1994; Karimi et al., 2014a). The %SS can be used as a separate measure or embedded in more detailed assessments, such as the Stuttering Severity Instrument (SSI) (see Chapter 5 for more details) for initial assessment, monitoring progress during a therapy programme, or assessing the effectiveness of different treatment programmes such as the Camperdown Program (O'Brian et al., 2010), Parent-Child Interaction Therapy (PCIT) (Kelman & Nicholas, 2017), and the Lidcombe Program (LP) (Packman et al., 2014). Although the LP has modified %SS to be an optional outcome measure (Bridgman et al., 2011), %SS offers a simple calculation involving the total number of syllables and the number of stuttering events occurring in a given speech sample. It is calculated using the following formula:

%SS = (total number of stuttered syllables / total number of syllables) ×100

However, although %SS may be regarded as a gold standard measure, it is prone to shortcomings that threaten its reliability and validity. One issue around this measure is the lack of consensus on how stuttering is measured. As mentioned in the previous chapter, defining stuttering is not a straightforward task. Many definitions have been offered, and there is no clear preference for one particular definition over another. In addition, there is little consensus as to what constitutes a stuttering moment and what does not. For example, there has been extensive debate on whether or not to consider whole-word repetitions as a symptom of stuttering (Howell, 2013; Howell & Davis, 2011; Jiang et al., 2012). Also, some researchers have advocated counting both stuttering moments and dysfluent behaviours (Wingate, 1964), whereas other researchers have recommended only counting stuttering moments (Costello et al., 1984). additionally, there is disagreement on whether to use

syllables or words as a unit of measurement, although, currently, counting syllables is more common than counting words. Another challenge affecting %SS is the day-to-day fluctuations in the stuttering levels of stutterers. As stuttering variability is well documented, the frequency of a speaker's stuttering, in addition to stuttering intensity and duration, differs markedly from situation to situation and from day to day (Constantino et al., 2016; Karimi et al., 2013). Therefore, it is difficult to obtain representative samples, especially since stuttering in some clients may improve in a clinical setting and thus fail to reflect their speech in everyday life (Throneburg & Yairi, 2001). Moreover, %SS only measures one aspect of stuttering and does not provide sufficient information to determine the severity of stuttering since it does not take into account the type of stuttering, duration of stuttering, or the physical tension associated with it, making the sole reliance on %SS problematic (Guitar, 2006). For example, with the %SS measure, mild stuttering may appear more severe than those who stutter less but a with longer duration, which can be perceived as more severe if evaluated perceptually.

Another issue with %SS is that, although precision in its counting is essential, its reliability has been questioned throughout the years. Reliability can be defined as how well the measure provides the same score when used several times in the same way (Guitar, 2006). To help ensure that the data is being gathered systematically and accurately, a reliability investigation may be carried out, for which two basic measures exist: inter-rater and intra-rater reliability (Ward, 2018). In this regard, inter-rater reliability refers to a situation where two or more raters apply the same measure on a sample on one occasion, whereas intra-rater reliability is where the same rater applies the same measure on the same sample on separate occasions (Karimi et al., 2014a). Historically, acceptable error is calculated via a percentage by the Pearson correlation. However, more robust statistical methods to assess agreement have been proposed (Karimi et al., 2014a) (for more details, see Section 7.4.2.2 in Chapter 7). The first study that can be traced to this issue is that of Tuthill (1940), involving 20 clinicians counting stuttering events, whose results showed 37–136 counts of stuttered moments in the same speech sample, reflecting the poor reliability of this measure. In a similar manner, some reports describing poor inter-rater reliability have been documented (e.g. Brundage et al., 2006; Cordes & Ingham, 1994, 1995, 1999; Ingham & Cordes, 1992; Karimi et al., 2014a; Kully & Boberg, 1988). In response to the issue of unsatisfactory reliability, different procedures have been employed and investigated. In this regard, the time-interval procedure has been used to rectify the low reliability associated with %SS (Alpermann et al., 2010,

2012; Cordes et al., 1992, 1999). This method requires a judge to identify the presence of any stuttering moment within a certain time interval (e.g. five seconds) and, instead of counting the number of syllables stuttered, the number of intervals that contain stuttering moments are counted. Intervals that are judged to have at least one stuttering event are labelled as stuttered intervals, whereas intervals that do not encompass any stuttering event are labelled as non-stuttered intervals (Cordes et al., 1992). The results of time-interval studies have indicated improved inter- and intra-judge agreement (Cordes et al., 1992). However, this method has been criticised, given that subdividing speech into intervals and then randomly presenting them to judges may influence the assessment of co-articulatory gestures and may therefore be misleading (Guntupalli, Kalinowski & Saltuklaroglu, 2006). In addition, the use of a time-interval methodology may lead to an under- or overcounting of stuttering events, as numerous stuttering events may happen in one interval and be regarded as one stuttering event, or single stuttering events may occur in different time intervals. Due to these issues, Guntupalli, Kalinowski and Saltuklaroglu (2006) concluded that, although this methodology may increase the reliability of counting stuttering, it also compromises its validity (Guntupalli, Kalinowski & Saltuklaroglu, 2006; Valente et al., 2015). The time-interval methodology has also been criticised as not being applicable in clinical settings in that it is not clear how it can aid clinical decision making; in other words, the time-interval methodology does not provide a frequency value with clinical significance (Guntupalli, Kalinowski & Saltuklaroglu, 2006; Valente et al., 2015; Yaruss, 1997).

Another approach adopted to improve the reliability of %SS is to listen to the sample twice, where the first time the judges count the total syllables and the second time the judges count the stuttered syllables (Jani, Huckvale & Howell, 2013; O'Brian et al., 2013). In this regard, both Jani, Huckvale and Howell (2013) and O'Brian et al. (2013) indicated the non-superiority of one method over the other. Also, in O'Brian et al. (2013), syllables were counted at a reduced speed when compared to the normal speed but yielded similar results and did not significantly differ from the standard approach of %SS counting.

On the other hand, several studies have indicated high levels of inter- and intra-rater agreement of %SS (Boberg & Kully, 1994; Ingham et al, 2001; Lincoln & Onslow, 1997; O'Brian et al., 2003, Onslow et al., 1996; Onslow et al., 1997). It was also reported that a high level of agreement is achievable when highly-trained judges take part (Brainbridge et al., 2015). However, there has been criticism that studies reporting high reliability are mostly

treatment studies that involve a low number of judges for their reliability checks (e.g. two judges) or studies that had small sample numbers (Valente et al., 2015). In short, evidence regarding the reliability of %SS is of a mixed nature. Nevertheless, it remains a central measure in evaluating treatment efficacy and monitoring change among participants.

A factor that has also been thought to influence %SS reliability is the presentation mode of speech samples (i.e. audio vs. video). When recording a speech sample for the purpose of calculating %SS, it is recommended to record a video speech sample rather than an audio speech sample because inaudible stuttering behaviour and physical concomitants may be missed in audio recordings (O'Brian et al., 2015). A comparison of %SS made from audio-visual and audio recordings was investigated in Rousseau et al. (2008), who compared the %SS ratings of preschool children who stutter using audio and audio-visual recordings and found that %SS was significantly lower, by 20%, when determined through audio recordings. In particular, this difference was thought to be attributed to the differences in the counts of stuttered syllables and not to differences in the total numbers of syllables spoken. In another study, four experienced SLTs measured %SS in ten-minute audio-visual and audio-only recordings of 36 adults who stutter, and there was a mean 18% increase in %SS scores when samples were presented in the audio-visual mode when compared with the audio-only mode (O'Brian et al., 2015).

Another factor that may contribute to the reliability of %SS is the speech sample size gathered. It has been suggested that larger sample sizes in different situations may be more representative and that the sample gathered should have an adequate number of target stuttering behaviours. However, investigations of the role of sample size on the disfluency data gathered are rather inconclusive, as one cannot identify a certain sample size as the gold standard (Sawyer & Yairi, 2006). Sawyer and Yairi (2006) used 1,200 syllables in their research with children who stutter, while, in other reports, there were nonsignificant differences in the number of stuttering disfluencies in different sample sizes, ranging from 300 to 1,800 syllables (Logan & Haj Tas, 2007; Robert, Meltzer & Wilding, 2009). Riley (2009) suggested 200 syllables in the SSI and Guitar (2014) recommended at least 300 syllables in spontaneous speech samples and 200 syllables for reading.

2.3.1.2 Speech rate

Speech rate is a measure that may reflect the severity of an individual's stuttering, and it is usually incorporated in conjunction with the stuttering severity measure (%SS) (Chon,

Sawyer & Ambrose, 2014). Generally, the speech rate may decrease for two possible reasons, either because the speech displays long prolongations, blocks, and many repetitions (i.e. severe forms of stuttering; Bloodstein, 1995: p.7) or as a result of applying a speech restructuring strategy, such as in prolonged speech during therapy (Teshima et al., 2010). The speech rate is typically calculated as either the number of syllables (syllables per minute [SPM]) or words per minute (WPM) in a given speech sample, and the syllable count is the unit of preference. When calculating SPM, all disfluencies are included in the total speaking time. However, extra repetitions of phrases, words, or syllables are excluded from the syllable count, as the final syllable total encompasses syllables that have conveyed meaningful information to the listener (Chon, Swayer & Ambrose, 2014). On the other hand, a related measure referred to as the articulatory rate is typically defined as the number of syllables or words spoken divided by the time required to produce the speech sample, while disfluencies, pauses, silent intervals, and so on are excluded (Hall, Amir & Yairi, 1999). However, SPM has been challenged as a valid and reliable measure, given that it suffers from some drawbacks. In this sense, it is a difficult task to distinguish normal pauses during speech with pauses that are the result of blocks and breaks in speaking, in addition to there being a lack of agreement as to what is considered typical in terms of speech rate (Ward, 2018).

2.3.1.3 Speech Naturalness

It can be argued that speech naturalness (NAT) is not a stuttering measurement *per se*; however, it is an important aspect to keep in mind when treating older children, adolescents, and adults who stutter. It is well documented that applying speech restructuring treatments, which are effective in reducing stuttering, may result in one's speech sounding unnatural (Ingham & Onslow, 1985; Ingham & Packman, 1977; Perkins et al., 1974). Therefore, it is essential to include NAT measures while applying treatment to ensure that, along with stuttering reduction, the client's speech sounds as natural as possible (Ingham & Onslow, 1985). A widely used naturalness scale to document speech naturalness is the 9-point scale developed by Martin, Haroldson and Triden (1984), which has been shown to be reliable, is user friendly, and requires no training (Armson & Kiefte, 2008; Hargrave et al., 1994; Ingham, Gow & Costello, 1985; Onslow, Adams & Ingham, 1992; Stuart & Kalinowski, 2004; Van Borsel & Eeckhout, 2008); this scale is now commonly applied in the Camperdown Program (see Section 2.4.3), where clients learn to rate the naturalness of their restructured speech using the nine-point NAT scale (O'Brian, Packman & Onslow, 2008).

2.3.1.4 Stuttering Severity Instrument (SSI)

See Chapter 5

2.3.1.5 Severity Rating (SR) scales

The Severity Rating (SR) scale is a perceptual measure that is simple to use and requires no equipment (Onslow, 2017), sometimes referred to as an equal-interval ordinal scale (Onslow, 2017). This scale is, by concept, different from %SS, where the rater is not required to count stuttering moments but rather assign a numerical value that represents perceived overall stuttering severity (O'Brian et al., 2004). When assigning a number to the sample heard, the rater takes into consideration the stuttering rate, stuttering frequency, and the severity of individual moments of stuttering (O'Brian et al., 2004). The assignment of a numerical value is usually done by the SLT. However, in some cases, the client can self-assign an SR score (for example, parents are required to report the SR scores of their children in the Lidcombe Program). The assignment of a numerical value of SR is of an arbitrary nature, and the severity rating scales differ according to the number of scale divisions they have, although commonly used are 7-point, 9-point, and 10-point scales (Onslow, 2017). For example, a 7-point severity rating scale has been used in the Illinois Early Childhood Stuttering Project, where 0=normal disfluency, 1=very mild stuttering, and 7=very severe stuttering (Yairi & Ambrose, 1992). Another example common in research is an SR scale from 1–9, where 1=no stuttering, 2=extremely mild stuttering, and 9=extremely severe stuttering (O'Brian et al., 2004), as, for example, used in the Camperdown Program (O'Brian et al., 2010). It is worth noting that using different scales incorporating different numerical values does not yield different results (Cullinan, Prather & Williams, 1963; Curran & Hood, 1977). As mentioned earlier, SR scales are simple to use and require no equipment and little or no training to apply. In addition, the reliability of SR appears not to be influenced by the level of training, as both experienced and inexperienced judges produced reliable results (Cullinan, Prather & Williams, 1963; Curran & Hood, 1977; Eve et al., 1995; Lewis & Sherman, 1951; Sherman, 1955; Young, 1969). High levels of inter-rater agreement have also been reported for severity rating scales when used with young children (Yairi & Ambrose, 1999; Yairi, Ambrose & Niermann, 1993).

Moreover, unlike %SS, SR scales are not affected by the mode of presentation (i.e. audio-only/audio-visual), as it has been shown that judges' scores on a 9-point scale did not differ across the two presentation modes (Rousseau et al., 2008).

The superiority of SR scales compared to %SS with regard to reliability may be logical, as it has fewer potential scores when compared to %SS, which requires counting events (Onslow, 2017); thus, the potential benefits of using a global scale like the SR instead of %SS gave rise to different investigations aiming to understand how interchangeable these scales are and whether or not the use of the SR instead of %SS is a valid option.

In a study by O'Brian et al. (2004), a correlation of .91 was reported for clinician ratings of %SS and clinician SRs incorporating a 9-point scale for 90 adults. For this reason, the authors suggested that %SS and SR are largely interchangeable. However, there are reservations to this conclusion. In cases where there are a small number of blocks and prolongations or a large number of repetitions within a speech sample, it is not sufficient to rely on SR alone; instead, a combination of %SS and SR is recommended. In a recent study by Karimi et al. (2014), 87 people who stutter received a ten-minute unscheduled telephone call, three SLTs measured these telephone samples using %SS and SR (9-point scale), and the results revealed a high correlation between the two measures. In another study by Karimi et al. (2014a), the authors evaluated %SS and SR reliability with more robust statistics than previous studies, as they combined the use of intraclass correlation coefficients (ICC) and limits of agreement (LOA), measures which enabled the authors to evaluate both relative and absolute reliability. Relative reliability aims to distinguish between clients, and absolute reliability, or what is referred to as agreement, aims to distinguish within subject changes (see Chapter 5 for more details). While the results revealed satisfactory relative reliability, absolute reliability was poor for both %SS and SR. Therefore, the authors concluded that %SS and SR are appropriate for use in research applications for comparisons of groups of participants, but, given the poor absolute reliability, these measures are not appropriate for evaluating small changes within individual participants. Interestingly, a study by Onslow et al. (2018) investigated whether parent-reported SRs provided similar estimates of effect size as %SS for randomised trials of LP. In this study, the dataset encompassed three randomised controlled trials of the Lidcombe Program (Arnott et al. 2014; Bridgman et al., 2016; De Sonneville-Koedoot et al., 2015) and data was analyzed using median changes and 95% confidence intervals per treatment group, Bland–Altman plots, analysis of covariance, and Spearman rho correlations. The results revealed no advantage of using %SS over parent-reported stuttering SRs as a primary outcome measure for the clinical trials chosen, and the authors suggested that a parent-reported rating of the child's typical stuttering severity for the week or month prior to each assessment is a justifiable alternative to %SS as a primary outcome measure in clinical trials

of LP, in addition to stating that it may accelerate the publication rate of clinical trials for that clinical population, given the practicality of parent-reported SRs. However, the reliance on severity rating scales has been criticised, as studies have shown that the use of the equal interval scale is not as accurate as the use of direct magnitude estimation (Schiavetti et al., 1983). For example, scores of 1–3 are likely to be more closely spaced than higher scores using a 1–9 scale (Riley, Riley & Maguire, 2004).

2.3.2 Attitudinal and Quality of Life (QoL) scales

Unlike the stuttering measures discussed earlier, which are concerned with the measurement of overt aspects of stuttering, QoL and attitudinal scales focus on the cognitive and affective aspects of the stutter; in this sense, attitudinal scales are designed to understand the unseen aspects of stuttering, such as feelings, attitudes, and perceptions, whereas QoL measures are concerned with the impact of stuttering on the speaker's quality of life, with such measures usually taking the form of self-report scales. It is not an easy task to identify core attitudinal and QoL measures, mainly because many such scales are available and comparatively few are used consistently in clinical practice and research (Susca, 2006). Thereby, the goal of this section is to discuss the scales that are commonly used in clinical contexts and have recently been incorporated in clinical trials of stuttering.

2.3.2.1 Communication Attitude Test (CAT)

The CAT is a self-report questionnaire designed to assess the speech-associated beliefs of school-age children who stutter. It has also been used successfully to assess children with other speech disorders (De Nil & Brutten, 1990). The CAT was developed by Brutten (1984) and was initially a 35-item questionnaire. However, it went through several modifications, and its latest version consists of 33 statements related to talking, e.g. 'I don't talk right' or 'My classmates don't think I talk funny', that the child classifies as either 'true' or 'false'. One point is given for each answer indicating a negative attitude toward communication so that a high score indicates an overall negative attitude towards communication whereas a low score indicates a positive attitude (Johannisson et al., 2009). The CAT is part of the Behaviour Assessment Battery (BAB), also consisting of the Speech Situation Checklist, which evaluates emotional reactions to speech disruptions in different speech situations, and the Behavioural Checklist (BCL), which identifies coping responses used by CWS to deal with their stuttering. Psychometric investigations of, and normative data from, the CAT suggest that it has satisfactory internal consistency (Brutten & Dunham, 1989; Brutten,

Vanryckegn & Vanryckeghem, 2007; Johannisson et al., 2009) and good test-retest reliability (Vanryckeghem, De Niels & Vanrobaeys, 2015) and validity (Brutten & Vanryckeghem, 2007; Johannisson et al., 2009). In addition, the Communication Attitude Test for Preschool and Kindergarten Children Who Stutter (KiddyCAT), a version for younger children, specifically preschoolers, has been developed (Vanryckeghem & Brutten, 2007), the aim of which is to explore the cognitive aspect of stuttering in preschoolers by administering a 12-item questionnaire. As in the CAT, the KiddyCAT consists of statements that investigate the child's beliefs about his/her speech, with the child being asked to respond to the statements with 'yes' or 'no', e.g., 'Do mom and dad think that you speak well?' Research carried out in the United States indicated that the KiddyCAT was able to differentiate between children who do not stutter and CWS for ages three–six (Vanryckeghem, Brutten & Hernandez, 2005), indicating that negative speech attitudes were reported for CWS as young as three years old, and these were significantly more negative when compared to non-stuttering children from the same age group (Vanryckeghem, De Niels & Vanrobaeys, 2015).

2.3.2.2 The Overall Assessment of the Speaker's Experience of Stuttering (OASES)

The Overall Assessment of the Speaker's Experience of Stuttering (OASES) is a relatively new questionnaire; it is patient-administered, designed to provide a comprehensive assessment of 'the experience of the stuttering disorder from the perspective of individuals who stutter' (Yaruss & Quesal, 2006), and is based on the World Health Organization's International Classification of Functioning, Disability and Health (ICF) (Yaruss & Quesal, 2006). This classification permits the description of disorders using three dimensions: impairment in body function or structure, contextual factors, and limitations or restrictions. In the context of stuttering, impairment in body function refers to the observable features of stuttering; contextual factors can be of a personal (e.g. avoidance, shame, and low confidence) or environmental nature (e.g. reactions by people in the speaker's environment); and limitations or restrictions refer to the impact of stuttering on daily life (Yaruss, 2010). The questionnaire has 100 items distributed among four categories: general information, reactions to stuttering, communication in daily situations, and quality of life, with the questions utilising a 5-point Likert scale. There are three existing versions for different ages: the OASES-S for school-age children 7–12 years old, the OASES-T for adolescents 13–17 years old, and the OASES-A for adults 18 years old and over.

The OASES has been commonly utilised in stuttering clinical research to measure changes over time (Franic & Bothe, 2005; Yaruss & Quesal, 2006), and it is usually used to complement behavioural stuttering measures that are clinician-led in order to offer a description of stuttering from the speaker's perspective. However, its use as an outcome measure has been questioned. Specifically, it has been claimed that, although the OASES items are based on the World Health Organization's definition of health, the number of items under each section of the ICF model is somewhat arbitrary (Franic & Bothe, 2008), thus challenging the representativeness of its items. For example, both the OASES and the Wright and Ayre Stuttering Self-Rating Profile (WASSP) (see Section 2.3.2.3 for more information on the latter) are based on the ICF model; however, the number of items under the impairment section are different in these two assessments. Although these two measures have different scorings, this example illustrates the nature of the arbitrary weighting of items in multi-dimensional assessments. It should be noted that the validation of the OASES is still in progress; for example, the factor structure of the OASES' subsections still warrants further examination through employing larger samples, as does the evaluation of the relationship between OASES results and participant demographics (Franic & Bothe, 2008). The OASES has been translated into different languages, and normative data is available for American, Dutch, and Australian populations. Arabic norms will be available soon (Yaruss, 2017, personal communication).

2.3.2.3 Wright and Ayre Stuttering Self-Rating Profile (WASSP)

The Wright and Ayre Stuttering Self-Rating Profile (WASSP) is a self-administered tool that attempts to describe aspects of stuttering in a comprehensive manner (Ayre & Wright, 2000). It was developed with the aim of measuring change over time and to help in setting clinical goals (Ayre & Wright, 2009). WASSP targets the overt, covert, and social aspects of stuttering, comprises 24 items distributed among five subscales and two additional optional items, and takes less than ten minutes to complete. The first subscale deals with the client's self-perception of overt stuttering (eight items), the second deals with the client's thoughts about stuttering (three items), the third with the client's feelings about stuttering (five items), the fourth with avoidance of stuttering (four items), and the fifth addresses self-perception of one's disadvantages due to stuttering (four items), with each question answered using a 7-point Likert scale (1=none; 7=very severe). As in the OASES, the WASSP is constructed based on the ICF model, and psychometric evaluation revealed satisfactory content and

criterion validity (Ayre 2009; Franics & Bothe, 2008). However, a thorough validity check is still warranted, despite its mainstream use both clinically and in research, especially in the UK (Ayre, 2009). The WASSP is designed for use with adults who stutter (18 years old and over), and a version for adolescents has also been developed (A-WASSP) (Ayre 2009). Further, the WASSP has been adapted as a visual analogue scale for children (Fibiger, Frøkiær & Jensen, 2007) and has been used by mothers to rate adolescents who stutter (Sønsterud, Mørk & Lind, 2007).

2.4 Stuttering Therapies

As in the case of stuttering assessment methods, stuttering treatment is a contentious matter in the field of speech and language therapy (Sidavi & Fabus, 2010; Yaruss, Coleman & Hammer, 2006). For example, in the case of preschool children, when is the optimal time to commence treatment given the chance of natural recovery? (Donaghy & Smith, 2016). However, it is becoming clear that early intervention for the management of stuttering is essential. Typically, early stuttering intervention is led by the parents, with the clinician acting as a facilitator (Donaghy & Smith, 2016). Stuttering intervention during the preschool years can be divided into two treatment approaches, the first of which is direct treatment, a type of stuttering intervention that aims to reduce stuttering and maintain stutter-free speech. From a theoretical point of view, stuttering reduction happens as a consequence of applying operant methods, motoric practice of stutter-free speech, or both (Donaghy & Smith, 2016), with a well-known example of such an approach being the LP (Onslow, Packman & Harrison, 2003). The second approach is multi-factorial treatment (or indirect treatment); unlike direct intervention, multi-factorial interventions do not target stuttering behaviours directly but, in order to reduce stuttering, rather target multiple factors within the child's environment that may trigger stuttering (Donaghy & Smith, 2016; Yaruss, 2002). This model derives from the Demands and Capacities Model (Adams, 1990; Starkweather & Gottwald, 1990), which is based on the idea that the onset of early stuttering is due to the limited capacity to speak fluently resulting from environmental, linguistic, emotional, and/or cognitive demands (Starkweather & Gottwald, 1990). Therefore, this type of intervention aims to change the behaviours of parents and family routines so as to decrease these demands (Donaghy & Smith, 2016), with a well-known example of such therapy being the Palin Parent-Child Interaction Therapy (PCIT; Kelman & Nicholas, 2017).

If stuttering persists after the preschool years, particularly from early school years to adolescence, its management requires consideration of multiple factors, including the severity of stuttering, stuttering tractability, and the developmental stage of the child (Donaghy & Smith, 2016). In the case of school-age children in their early years, therapy options, such as the adaptation of preschool stuttering treatments, have been investigated (e.g. Lidcombe). For adolescents who stutter, however, adult treatments (such as the Camperdown Program) have been adapted. In this regard, studies that have addressed the treatment of stuttering in adolescents have adopted different speech restructuring techniques, which may encompass practicing new speech patterns to replace stuttered speech, while sounding as natural as possible (e.g. prolonged speech) (Cocomazzo et al., 2012) or the adoption of techniques to change the delivery of speech to overcome moments of stuttering, known as fluency shaping techniques (Donaghy & Smith, 2016). Although the focus of this thesis is on assessment rather than treatment, it was felt that it is important to include a section discussing treatment options for CWS, as assessment and treatment are inseparable where stuttering measures facilitate therapy goal setting, give an indication whether there is progress towards treatment goals and help monitoring maintenance of therapy. In the following section, the three aforementioned programmes (Lidcombe, PCIT, and Camperdown) will be reviewed. It is acknowledged that there are more therapy approaches and versions of stuttering modification which are not covered in this section (Andrews et al., 2012; Boberg & Kully, 1985, 1994; Cooper & Cooper, 1985; Craig et al., 1996; Druce, Debney & Byrt, 1997; Hancock et al., 1998; Howie, Tanner & Andrews, 1981; McGuire, 2003; Menzies et al., 2008; Ward, 1992; Zebrowski & Arenas, 2011). These three therapy programs were chosen as they are arguably the three most commonly known approaches to the management of stuttering up to adolescent years, which encompass the age group of the main study (6-15 years) and are commonly applied in Saudi clinical settings.

2.4.1 The Lidcombe Program (LP)

The Lidcombe Program (LP) is an operant, parent-conducted treatment designed for preschool children (Onslow, Packman & Harrison, 2003) consisting of two stages. In Stage 1, the child and parents are scheduled to see the SLT for one hour on a weekly basis, where the SLT teaches the parents how to control the child's stuttering by utilising verbal contingent stimulation. The parents are taught to present three verbal contingencies for the child's stutter-free speech and two verbal contingencies for stuttering. During the course of

treatment, the SLT measures the child's stuttering frequency (%SS) weekly. In addition, the parents are required to measure the child's stuttering severity daily using a 10-point scale (1=no stuttering, 2=extremely mild stuttering, 10=extremely severe stuttering) (see Section 2.3.1.2). The child completes Stage 1 after he/she meets the treatment target, which is below 1% syllables stuttered and the parent severity rating of 1 or 2 for three consecutive weeks. After this, the child enters Stage 2, referred to as the maintenance stage, as its aim is to sustain the no stuttering or almost no stuttering that has been achieved in Stage 1 (Onslow, Packman & Harrison, 2003).

The developers of the programme state that, if correctly administered in adherence to its guidelines, especially with regard to the verbal contingencies, a complete and permanent recovery from stuttering is achievable (Onslow, 2017). Randomised controlled clinical trials revealed favourable outcomes for LP (Harris et al., 2002; Jones et al., 2005; Lattermann, Euler & Neumann, 2008; Lewis et al., 2008). In a recent meta-analysis of LP outcomes (Onslow et al., 2012), it was revealed that children who received LP were 7.5 times more likely to have stuttering below 1%SS at a later time than children who did not receive the treatment (Onslow & Millard, 2012).

Several treatment reports of the Lidcombe Program in Western and non-Western countries have been published, for example, in Australia (Onslow, Andrews & Lincoln, 1994), Canada (Lattermann, Shenker & Thordardottir, 2005), the Netherlands (Franken, Kielstra-Van der Schalka & Boelens, 2005), Germany (Lattermann, Euler & Neumann, 2008), New Zealand (Jones et al., 2005), Kuwait (Al-Khaledi et al., 2017), and Iran (Bakhtiar & Packman, 2009). As mentioned earlier, while LP was specifically designed for preschool children who stutter, its success with school-age children who stutter has also been reported (Rousseau, Onslow & Packman, 2005).

2.4.2 The Palin Parent-Child Interaction Therapy (PCIT)

The Palin Parent-Child Interaction Therapy (PCIT) is a treatment programme developed by the Michael Palin Centre for Stammering Children in London and is based on the multi-factorial model to explain stuttering, the theoretical model for which is that stuttering is triggered and sustained due to predisposed factors, such as motor, physiological, linguistic, and developmental factors (Kelman & Nicholas, 2017).

At the start of PCIT, a detailed consultation assessment is carried out, including the assessment of language, fluency, and communication skills, and an interview with the parents and the child is conducted to understand how aware the child is of the stuttering, the child's perspectives of the stuttering, and the impact it has on the child (Millard, Nicholas & Cook, 2008). In addition, the parents are interviewed for a detailed case history. This information is then gathered together in order to understand the physiological, linguistic, emotional, and environmental factors supporting or impacting on the child's fluency (Millard, Nicholas & Cook, 2008). It should be noted that PCIT is individualised for each child based on linguistic, environmental, and emotional strengths and needs. The assessment is followed by one session per week over six weeks of clinic-based therapy, and then a six-week period of home consolidation, during which the parents continue to implement the strategies and skills they have developed during the clinic therapy phase (Onslow & Millard, 2012). Efficacy studies of PCIT, particularly replicating single subject clinical trials in different settings, have been carried out, e.g. PCIT when delivered by specialist therapists in a tertiary specialist centre (Millard, Edwards & Cook, 2009; Millard, Nicholas & Cook, 2008). Moreover, studies have explored the treatment when delivered by therapists working in different clinical settings (Crichton-Smith, Baker & Rowley 2003; Matthews, Williams & Pring, 1997). In order to ensure that the reduction of stuttering is not due to natural recovery, all of the participants chosen for these studies had been stuttering for more than 12 months, (Onslow & Millard, 2012). Unlike the LP, PCIT does not aim to reach a zero or near zero stuttering level but rather a decreasing trend in stuttering, reduced parental anxiety, and increased parental confidence in managing stuttering (Onslow & Millard, 2012). Although the aims of this intervention are comprehensive, changes in stuttering frequency (%SS) is the primary outcome of PCIT.

2.4.3 The Camperdown Program (CP)

The Camperdown Program, developed by O'Brian et al. (2003) at the Australian Research Centre, is a speech restructuring programme wherein the primary aim is to reduce stuttering in everyday speech situations significantly. It also aims to help clients to adopt self-managed procedures so that they are able to deal with any increase in stuttering at any time. It should be noted that this programme does not recommend any specific techniques to address the social anxiety related to stuttering. However, it offers flexibility during the problem-solving sessions with regard to adopting such procedures when needed (O'Brian et al., 2010). The CP

can be carried out in an intensive format over one week or it can be conducted on weekly clinic visits. In addition, the CP can either be implemented in a group therapy format or in one-to-one sessions. CP has also been adapted to be carried out face-to-face or through using the telephone (telepractice). The CP was originally a therapy programme designed for adults who stutter. However, it has also been trialled on children between the ages of 12 to 17, with the results indicating an overall reduction in stuttering severity. However, in terms of the efficiency of CP, the effect for this age group has been inconsistent (Carey et al., 2014; Carey et al., 2012; Hearne et al., 2008).

3 Availability and Access to Speech and Language Therapy Services in Saudi Arabia

3.1 Introduction

The need for stuttering management services for school-age children is essential, given the diverse consequences of stuttering for that age group, as children who stutter in school environments are more susceptible to bullying, along with the challenges posed for social participation (Davis, Howell & Cooke, 2002; Daniels, Gabel & Hughes, 2012) (See Section 1.6 in Chapter 1 for more details on the negative consequences of stuttering). Although the chances of recovery for CWS entering their school years decrease (Yairi & Seery, 2015), and stuttering therapy focuses on management rather than prevention (Yaruss, Coleman & Quesal, 2012), it is critical to learn effective management techniques before the disorder becomes an entrenched and lifelong pattern (Nippold, 2011). The accurate measurement and assessment of stuttering is key to the effective management of the disorder (for more details on the importance of stuttering measurements, see Section 2.2 in Chapter 2). For many in Saudi Arabia, however, this is currently unachievable due to a range of access barriers and service limitations preventing children from receiving assistance from specialised SLTs. As mentioned earlier, given the vast evidence in the literature of the challenges faced by individuals who stutter throughout their lives, this thesis investigates a model of service delivery that can increase access to SLT stuttering services, particularly stuttering assessment.

In this chapter, an introduction to Saudi Arabia's demographics is provided, along with a discussion of the availability and access to SLT services, in order to pinpoint how difficult it is for some children and their families to access services in a timely and effective manner.

3.2 Saudi Arabia

Saudi Arabia, covering over 2,150,000 square kilometres (830,000 square miles), is the largest country in the Middle East, with a population of around 28,376,355 people (Central Department of Statistics and Information, 2011). Saudi Arabia is divided into 13 administrative regions (see Figure 1), but it can be culturally divided into five distinct zones: Northern, Southern, Eastern, Western, and Central Saudi Arabia. Most of the population (65.0%) is concentrated in three main administrative regions: Riyadh (Central), Makkah (Western), and the Eastern region (Central Department of Statistics and Information, 2011). A

remarkable feature of Saudi Arabia's population is that it is dominated by a young demographic, where people under 20 form 49.3% of the total population (UNICEF, 2013).

Figure 1 Map of Saudi Arabia

3.3 Availability and Access to SLT Services

Despite the importance of delivering speech and language therapy services to CWS (especially given stuttering's sometimes lifelong adverse effects), a study by Wilson, Lincoln and Onslow (2002) in Australia showed that there can be numerous barriers to accessing speech and language therapy. In this sense, Wilson et al. (2002) interviewed speech and language therapists who provided services to children in rural areas, and their data suggests that speech therapy services are unequally offered across Australia, where paediatric patients living in rural areas face challenges in access or receive therapy that is compromised in

quality. Barriers to accessing speech and language therapy services include long distances, long waiting lists, the cost of providing services, the lack of public transportation, and a lack of awareness of available services (O'Callaghan et al., 2005; ASHA, 1985). It is worth noting that limited access to speech and language therapy services is not exclusive to people living in rural areas (Waite, 2010). Regular attendance at speech and language therapy clinics may also be a challenge for parents who work full time, have large families, or are facing financial difficulties. Therefore, increasing demands for SLT services influence both metropolitan and rural areas (Waite, 2010). Barriers to accessing SLT services in metropolitan areas include long waiting lists (Victorian Department of Human Services, 2004) and a difficulty in recruiting experienced staff (Iacano et al., 2007). In essence, a shortage of speech and language therapists exists in both the health and education sectors in metropolitan, suburban, and rural areas in the United States and Australia (ASHA, 2008; Waite, 2010).

In Saudi Arabia, despite the country's relatively high expenditure on healthcare, where investment in health services by the Ministry of Health (MOH) witnessed a rise from 6.49% of total government expenditure in 2010 to 7.27% in 2015 (Ministry of Health, 2015), access to SLT services is not much different than that in the United States and Australia. Addressing the public's inequitable access to resources is considered a major issue in Saudi Arabia's healthcare system (Alkabba et al., 2012), where regular provisions, including SLT services, are not available to many of the country's children living in either major cities or rural or remote areas.

In a recent study by Khoja and Sheeshah (2018), a telephone survey was conducted to explore the availability of SLT services in the Saudi public health sector, with data revealing that SLT services are severely lacking, which was attributed to a shortage of SLTs and the unequal distribution of facilities among major cities. In addition, several barriers to SLT service development were identified, including limited numbers of SLTs, with the ratio of SLTs to inhabitants in Saudi Arabia being .67 per 100,000 (in contrast, the UK has a ratio of 16.3 per 100,000) (Fagan & Jacobs, 2009). Although the lack of SLTs is a major problem, financial concerns are not an issue in Saudi Arabia, as most services are well equipped (Khoja & Sheeshah, 2018). However, it was reported that many hospitals featuring MOH rehabilitation services encompass physical therapy for outpatients with less emphasis on SLT, which can be attributed to a lack of awareness of the SLT profession. This stems from the fact that SLT is relatively new in the region, in addition to a lack of research, assessment tools, and

interventions designed specifically for the Arabic population (Al-Jadid, 2013; Khoja & Sheeshah, 2018; Shaalan, 2009).

As mentioned earlier, Fagan and Jacobs' (2009) study revealed an unequal distribution of SLTs in Saudi Arabia, where most services are centralised in the largest cities and most rural areas do not receive local speech and language therapy services, being either underserved or not served at all. Moreover, there is a significant shortage of SLT services even within major cities, where the field of communication disorders is still emerging and speech and language therapy treatment facilities are not widely available. For example, only healthcare settings typically provide SLT care, whereas such services in schools are severely lacking (Alquraini, 2010).

This conflicts with the World Health Organization's 'Health for All in the 21st Century' strategy (WHO, 1998), the objective of which is to make quality healthcare available to all, as equity in healthcare for each country and its population is listed as a basic value of the WHO's current policy framework (WHO, 1998).

The primary reason for such poor performance may be due to the fact that Saudi Arabia is a large country, wherein people must travel long distances to get from one populated area to another (Saudi Arabia Ministry of Health, 2010; Saudi Arabia Ministry of Health, 2009; WHO, 2008). For example, travelling between Wadi Al-Dwaser and the centre of Riyadh, where SLT services are available, is a seven-hour drive (740km), highlighting the severe lack of overall SLT accessibility (Khoja & Sheeshah, 2018). In a study by Verdon et al. (2011), it was revealed that patients become hesitant to travel more than 50 km to receive SLT services.

Two studies conducted in Saudi Arabian hospitals showed high rates of missed health appointments, with hospitals in the eastern provinces having a 23.7% rate of missed appointments and those in the central provinces having a rate of 30% (El-Din, Al-Shakhs & Al-Oudah, 2008; Griffiths, 2001). The researchers attribute these high percentages to the amount of time that passes between scheduling an appointment and the actual date, in addition to the fact that public transportation is not available. Transportation has traditionally been an even greater challenge for Saudi Arabian women, as they are not allowed to drive (El-Din, Al-Shakhs & Al-Oudah, 2008; Mohamed & Al-Doghaither, 2002), although this driving ban was suspended in 2018.

With regard to SLT services for people who stutter in particular, given the lack of SLTs and the unequal distribution of SLTs in Saudi Arabia, specialised SLT services including stuttering cannot be recognised, as most clinical facilities offer general SLT services that are mainly found in medical governmental settings or in private practice. This may be problematic, as surveys show that general clinicians believe their skills are not sufficient to assess and treat stuttering due to limited training (Kelly et al., 1997; Tellis, 2007).

In summary, SLT services in Saudi Arabia suffer from multiple shortcomings, including a deficit of SLTs and an unequal distribution of those that do exist, coupled with general ignorance about their presence and the services they provide. As for SLT resources, there is a lack of evidence-based practice in the field of stuttering in Arabic language and culture, and, to our knowledge, no standardised tests or interventions have been designed specifically for this particular population; rather, there is a forced reliance on programmes adapted from Western models. The exact nature of assessment and treatment is therefore not clear, as there are no specific guidelines to be followed.

4 Identifying the Assessment Practices of Saudi Speech and Language Therapists for School-Age Children Who Stutter

4.1 Introduction

Prior to conducting the main study of this thesis, given that the field of speech and language disorders is still emerging, it is essential to understand the assessment practices followed by SLTs in Saudi Arabia, particularly their assessment practices in the field of stuttering, in order to help in the selection of the most relevant assessment activities related to stuttering in Saudi clinical settings. Once these common procedures are identified, the formal investigations described in Chapter 7 will be carried out.

4.2 Questionnaire development

The questionnaire was drafted after consulting SLTs in the UK and Saudi Arabia in the field of stuttering and conducting a literature review. The literature review included searching articles and textbooks on stuttering assessment and its related issues. After constructing the questionnaire, it was distributed to several SLTs in Saudi Arabia, who piloted it. The questionnaire was then modified based on their feedback.

In order to gain specific information about SLTs, the final version of the questionnaire (see Appendix 1) consisted of 26 questions divided into three sections: multiple-choice, fill-in-the-blank, forced-choice questions, and open-ended questions. The first section was designed to collect background information relating to age, gender, qualifications, years of experience with stuttering, work setting, and typical caseload. In the second section, SLTs were required to answer questions regarding the typical assessment procedures used with school-age children who stutter. Information was gathered on the use of formal and informal tests, additional tests, stuttering assessments used, case history forms used, specific components included in a stuttering assessment, whether attitudes and quality of life is addressed in the assessment, and time taken for an assessment session. The third section included questions about the patients' locations, how often SLTs see patients outside their cities, how often patients fails to attend due to their distant locations, and whether performance would differ if an SLT was locally accessible. There were also three questions assessing SLTs' familiarity with and willingness to include telepractice in their services.

4.3 Participants and procedure

Ethical approval from the University of Reading was given for this investigation on 3 March 2015. The questionnaire was distributed using both paper-based and web-based methods, using the Saudi Society of Speech and Language Pathology and Audiology Database. For paper-based formats, an assigned individual in Riyadh was asked to manually distribute the questionnaires and then collect them within one week. After collection, the individual was asked to scan the questionnaires and email them to the doctoral researcher. However, due to the difficulties in applying this method (e.g. poor return rate and limited distribution), a web-based questionnaire using 'Qualtrics' online surveys was applied to reach a larger number of SLTs throughout Saudi Arabia. The questionnaire link was sent via email and mobile services, and SLTs who received the link were asked to forward the participation link to relevant colleagues in order to increase the response rate. The survey was open for eight months, from March to October 2015. Since SLTs working in Saudi Arabia usually have mixed caseloads (both paediatric and adult and multiple and diverse aetiologies), the survey was distributed to all SLTs to ensure that it was received by a large number of SLTs working with CWS. Each SLT received the questionnaire and a cover letter explaining the purpose of the investigation. Due to the lack of an official registry, the exact number of Saudi-certified SLTs could not be obtained, and the exact response rate was difficult to calculate. Twenty questionnaires were returned, although only 16 were included in the final analysis, as the other 4 were incomplete.

4.4 Data analysis

Quantitative data was analysed using descriptive statistics through the calculation of frequencies. Since the small sample size hindered the use of inferential statistical analysis, the data is presented descriptively and must be interpreted with caution. In addition, the open-ended questions were analysed descriptively.

4.5 Results

4.5.1 Section 1: Background information (age, gender, qualifications, years of experience with stuttering, work setting, and typical caseload)

Eleven respondents (68.8%) were aged 21–29 and five (31.2%) were 30–39; two (14%) were male and fourteen (87.5%) were female; twelve respondents (75%) held a bachelor's degree, three had a master's degrees (18.8%), and one (6.3%) held a PhD.

All respondents were from Riyadh, and all had received their education at King Saud University, Riyadh (the only university that offered an SLT course at that time). Moreover, two respondents reported receiving higher education in the UK and one had studied in the United States. Respondent's years of experience as SLTs generally ranged from 1–15 years, with a mean of 5.1 years, while their years of experience with stuttering ranged from 1–15, with a mean of 3.5. Regarding their employment settings, two respondents (12.5%) reported that they mainly worked in in-patient acute hospitals, four (25%) worked in out-patient acute hospitals, three (18.8%) worked in in-patient rehabilitation settings, six (37.5%) worked in out-patient rehabilitation settings, and one (6.3%) was in private practice.

As for the typical caseload of stuttering patients, five respondents (31.3%) reported that 5–10% of their caseload consisted of patients who stuttered, five (31.3%) had from 10–25%, four (25%) had from 25–50%, one (6.3%) had from 50–70%, and one (6.3%) had a caseload of 75% stuttering patients (see Table 1).

Table 1 Summary of Respondents' Demographic Characteristics

		N=16
Characteristic	**N**	**%**
Age (years)		
21–29	11	68.8%
30–39	5	31.3%
Gender		
Male	2	12.5%
Female	14	87.5%
Highest Level of Education		
Bachelors	12	75%
Masters	3	18.8%
PhD	1	6.3%

Saudi Arabian City Riyadh	16	100%
Years of Experience in SLT Ranged from 1–15 years, with a mean of 5.1 *Years of Experience in Stuttering* Ranged from 1–15, with a mean of 3.5	-	-
Employment Setting In-patient acute rehabilitation hospital Out-patient acute hospital In-patient rehabilitation hospital Out-patient rehabilitation setting Private practice	 2 4 3 6 1	 12.5% 25% 18.8% 37.5% 6.3%
Typical Caseload of Stuttering Ranged from 5–10% Ranged from 10–25% Ranged from 25–50% Ranged from 50–70% 75%	 5 5 4 1 1	 31.3% 31.3% 25% 6.3% 6.3%

4.5.2 Section 2: Typical assessment procedures used with school-age CWS (the use of formal and informal tests, additional tests, stuttering assessments used, case history forms used, specific components included in a stuttering assessment, whether attitudes and quality of life is addressed in the assessment, and time taken for an assessment session).

Respondents indicated that the assessment of school-age children was conducted either through an SLT alone or in conjunction with a psychologist. Six (37.5%) respondents stated that they followed specific guidelines to assess CWS, whereas ten (62.5%) reported that they did not. The majority of those who used specific guidelines reported that guidelines involved the use of a detailed case history form and the application of the SSI-IV; fewer respondents reported that referrals and language screening were part of their guidelines. When asked about what types of assessments used (formal/informal), one (6.3%) relied on formal assessment, six (37.5%) reported their reliance on non-formal assessment, and nine (56.3%) reported the use of both. Respondents were asked to name the formal assessments used in their practice, and 14 (87.5%) reported using SSI-IV, whereas one (6.3%) used SSI-IV, OAESES, and WASSP. As to how often they adapted non-Arabic assessments, three (18.8%) said never, three (18.8%) said occasionally, and ten (62.5%) said frequently. Respondents were asked to name them, and 12 reported adapting the SSI-IV.

Moreover, they were asked about the specific components they usually incorporated in their assessment, and the majority indicated focusing on the overt aspects of stuttering, including %SS, types, and duration, as well as secondary behaviours, as part of their assessment. As for cognitive and emotional aspects, 13 (81.3%) reported assessing avoidance and emotional reactions to stuttering. In addition, respondents were asked if they assessed other aspects beyond stuttering, and 13 (81.3%) reported assessing language, 11 (68.8%) articulation, 5 (31.3%) phonology, and 1 (6.3%) reported incorporating cognitive screening for school-age CWS (for more details, see Table 2). When respondents asked whether attitudes and QoL were considered, 13 (81.3%) reported yes and 3 (18.8%) reported no. Those who answered yes were asked about how they tackled attitudes and QoL, and all respondents reported that they conducted an interview with the parents. Most reported that questions asked to parents related to the impact of stuttering on the child's social life and school performance; only three reported interviewing the child.

Respondents were also asked if, during their clinical practice, they used a specific case history form to assess CWS; half said they did and half said they did not. Those who responded positively were asked which questions they most commonly asked, the results of which were as follows: four asked about the age of onset, three asked about family history of stuttering, and six asked about awareness of stuttering. Less common questions included: cause of stuttering in the parent's opinion (two respondents), critical changes during stuttering emergence (two respondents), factors influencing the increase and decrease of stuttering (two respondents), patient reaction to stuttering at home and school (two respondents), avoidance (one respondent), and previous therapy (one respondent).

In addition, Respondents were asked if the case history form included direct questions to the parents; thirteen (81.3%) said yes and three (18.8%) said no. Those who reported yes were asked to name the top questions they addressed to the parents. One respondent asked about the possible cause of stuttering, two asked about awareness, four asked in what situations stuttering increased and/or decreased, one asked the age of onset, one asked about previous therapy, two asked about critical changes at home, two asked about reactions to stuttering, two asked about awareness, one asked about family relationships, one about performance at school, one about avoidance, one about sibling interactions, and one asked how parents reacted to their child's stuttering.

All respondents reported that assessments took one session, typically lasting for one hour.

Table 2 Assessment Components Included by Saudi SLT in a Stuttering Assessment

		N=16
Assessment Component	**N**	**%**
Overt		
%SS	16	100%
Types	16	100%
Duration	16	100%
Secondary Behaviours	16	100%
Rate (SPM)	8	50%
Effect of Task on Stuttering	11	68.8%
Effect of Situation on Stuttering	10	62.5%
Naturalness	4	25%
Covert		
Emotional Reactions	13	81.3%
Avoidance	13	81.3%
Expectation of Stuttering	7	43.8%
Expectation of Fluency	5	31.3%
Motivation	12	75%
Self-Perception	12	75%
Additional Components		
Language	13	81.3%
Articulation	11	68.8%
Phonology	5	31.3%
Cognitive Screening	1	6.3%

4.5.3 Section 3: Questions about the patient's location, how often SLTs see patients outside their cities, how often patients fail to attend due to their distant locations, and whether performance would differ if an SLT was locally accessible. There were also three questions assessing SLTs' familiarity with and willingness to include telepractice in their services.

When respondents were asked about the frequency of treating patients outside the city, one said once a year, five said once a month, and ten said every week.

Clinicians were asked to speculate about whether their patients' performance may have differed if services were available in their city: fifteen (93.8%) reported yes and one (6.3%) said no. When asked why, the respondents said that, if they were in the same city as the patient, then:

- the chance for intensive, regular therapy would increase and patients would make better progress.
- parents could attend the sessions and therefore become more compliant.
- commitment and motivation for therapy could increase.
- consistent training and application of learned strategies, which is a crucial component of improvement, would increase.
- continuous emotional support from the clinician would be available.

When asked how often patients failed to attend appointments, two (12.5%) said rarely, ten (62.5%) said occasionally, and four (25%) said frequently.

The reasons for missing appointments were:

- expenses
- other commitments (work and family)
- distance (when families live in rural areas, travelling may be inconvenient, especially if children need intensive therapy).
-

4.5.4 Section 4: Telepractice

When asked about telepractice, 11 (68.8%) respondents reported that they were not familiar with it (i.e. never heard of it), whereas 5 (31.3%) were familiar; four said they had used TP, while 12 had not; and twelve (81.3%) said they would consider using it, but three (18.7%) said they would not.

4.6 Discussion

The goal of the present investigation was to provide information regarding the stuttering assessment practices employed by SLTs in Saudi Arabia, with some questions addressing the SLTs familiarity with telepractice as a potential clinical vehicle as well as their willingness to incorporate it into their practice.

All respondents indicated that they assessed overt aspects of stuttering using various quantitative measures. All respondents reported the use of %SS, which is a vital component of monitoring the success of therapeutic approaches and programmes for school-age children who stutter, such as fluency-shaping programmes (e.g. prolonged speech and the Camperdown Program). However, only using %SS can be misleading when assessing the severity of stuttering; therefore, it is recommended to assess additional dimensions, such as the type and duration of stuttering or the physical tensions associated with it, which all respondents included in their assessments.

The majority of respondents also reported assessing cognitive and emotional issues relevant to stuttering, such as avoidance, emotional reactions, and expectations of fluency and dysfluency. These are in-line with current research, as there is increasing recognition of the value of the clients' point of view and the importance of assessing the behavioural, cognitive, and affective changes that occur as a consequence of stuttering (Guntupalli, Kalinowski & Saltuklaroglu, 2006; Ward, 2018; Yaruss, 2010).

When respondents were asked to name the formal assessments they used, the majority reported using SSI-IV, whose use is common in research and clinical contexts worldwide, with only one respondent reporting the use of OASES and WASSP, although such instruments are mostly used with adults. As mentioned earlier, the majority of the sample reported assessing cognitive and emotional issues relevant to stuttering. However, most respondents reported the use of interviews to assess QoL and attitudes rather than formal scales, and, although the exact reason for their reluctance to use such scales is not clear, it

may be due to the fact that, in Saudi Arabia, SLTs work across a range of areas in a generalist, rather than specialised, approach, dealing with both adult and paediatric populations across different communication disorders; therefore, they may not be aware of the most current tools available to assess various aspects of stuttering, or perhaps all the available scales are Western in origin and not standardised for the Arabic population. Therefore, SLTs may believe that conducting an interview with a client could be more informative. In addition, there may be a lack of resources; there are no authentically Arab-developed tests available for purchase. Even if they existed, such tools are usually unpublished tests that are part of PhD materials generated by Saudi students studying abroad, which are difficult to trace, in addition to clinicians not having direct access to them. Further, there are no publishing companies offering SLT general assessment tools or tools specific to stuttering in Saudi, which hinders the ability of SLTs to take advantage of such developed tools (Khoja, 2017).

An interesting finding is that, when SLTs were asked about how they assessed the attitudes and QoL of CWS, they answered that they achieved this by interviewing the parents only, the reason for which was not outlined in their answers; however, it may have been due to the outdated belief that SLTs should not make children aware of their stuttering behaviour, regardless of their age (Al-Khaledi et al., 2014). This misconception is related to Wendell Johnson's (1955) diagnosogenic theory, which states that stuttering is caused by the misdiagnosis of normal dysfluencies in a child's speech by the child's parents; that is, stuttering begins 'not in the child's mouth but in the listener's ear'. This theory gained popularity in the 1950s, but it is considered outdated nowadays (Al-Khaledi et al., 2014). These kinds of misconceptions may reflect the limited academic and clinical training in the field of fluency as, in Saudi Arabia, throughout the years of a BSc degree, only one fluency course is taught. After finishing the BSc, students are required to spend one year working in hospital settings as interns under clinical supervision, and treating children and adults who stutter during the internship year is dependent on the caseload of the supervising clinician. Therefore, there is a pressing need for encouraging SLTs to attend continuing education workshops to keep themselves updated about the latest assessment practices and treatment approaches.

Almost 62% of respondents reported that they provided SLT services to people who must travel to the clinic on a weekly basis and noted that their progress may be better if services were available locally. For example, they mentioned that the availability of local services

could lead to an increase in commitment to therapy, and there may be more emotional support from the clinician if services are available locally. However, as mentioned earlier in this chapter, SLT services are not readily available to children in need in Saudi Arabia, and therefore telepractice could be an option for increasing access to SLT services. However, prior to implementing the service delivery model, it must be checked for feasibility, validity, reliability, and acceptance in order to be implemented successfully, which is the aim of this thesis. In this survey, most respondents were unfamiliar with the TP delivery model. However, what is encouraging is that the majority (81.4%) reported that they would consider using it.

In summary, the questionnaire respondents assessed various dimensions of stuttering (both covert and overt). For the assessment of overt features of stuttering, most respondents reported the use of formal and quantifiable measures, such as the %SS and SSI; however, the reliability of the SSI has not been assessed in the Arabic context. Respondents reported using a non-formal approach for the assessment of attitudes and quality of life through interviews. This highlights the need for SLTs to receive more training in the domain of stuttering in order to make them aware of up-to-date assessment tools and teach them how to adapt well-established Western tools to the Arabic language or perhaps develop authentic assessment tools and scales for the target population.

Based on the findings of the questionnaire, it was decided that the SSI-IV would be used as an assessment task for this telepractice research. It was also decided that the %SS and SR would be included, as they are common measures in both clinical and research contexts. In addition, the reliability of the adapted SSI-IV will be investigated before implementation in the main study (see Chapter 5).

5 SSI: Background, Procedures, and Reliability

5.1 Introduction

Although Chapter 2 was devoted to discussing various stuttering measures, this chapter is dedicated solely to discussion of the Stuttering Severity Index (Fourth Edition) (SSI-IV), the reason for which is that it was decided to include the SSI-IV as a stuttering assessment task for the telepractice study, as the main focus of this thesis (see Chapter 8). This chapter a) describes the background and functioning of the assessment together with a discussion of its relative strengths and weaknesses, before concluding with a study outlining the processes and procedures involved in its translation into Arabic and a small pilot study that tests its relative and absolute reliability.

5.2 Background

The Stuttering Severity Instrument (SSI) is one of the most widely used syllable-based procedures for assessing stuttering symptoms (Riley, 1994; Riley, 2009). The SSI (Riley, 1972), currently in its fourth edition (SSI-IV), was developed as a means of clinically evaluating treatment efficacy and as a research tool for studying the effects of stuttering (Riley, 2009). In 1980, the SSI was modified, and then a further modification was carried out, resulting in the SSI-III, with the author pointing out that it only offers a partial description of stuttering severity and that diagnoses require attention to all the available information (Riley, 1994). The latest version, the SSI-IV, offers additional avenues of evaluation, including self-reports; samples beyond the clinic and telephone samples (Riley, 2009); the incorporation of the Computerized Scoring of the Stuttering Severity Version 2 (CSSS-2.0); the addition of a speech naturalness rating scale to the Examiner Record Form; and the inclusion of a scale to self-report feelings related to stuttering, the Clinical Use of Self-Reports. The SSI-III (Riley, 1994) and SSI-IV (Riley, 2009) were normed on a sample of 72 preschool-age children, 139 school-age children, and 60 adults (Riley, 2009).

The severity score obtained is based on three components of the instrument, which are derived from speech samples: 1) the percentage of syllables stuttered (%SS), 2) the duration of the longest three stutters within a sample, and 3) physical concomitants, which are based on physical behaviours related to the stutterer's speech. In the SSI, conversion tables are provided for converting the values for each component, after which they are summed to provide the total score of the SSI. In addition, these three measures are collected in different

ways based on whether the child can read or not. For more details on the procedures, see Appendix 2.

In terms of the design features of the SSI-IV, four strengths can be identified (Todd et al., 2014). First, events that are considered as stuttering moments and those that are not are clearly stated. This is explained in detail in the SSI-III manual, and the procedure is identical in SSI-IV. The following events are included as stutters: 'repetitions or prolongations of sounds or syllables (including silent prolongations)' (Riley, 1994: p.4). Riley (1994) also notes some of the events that are not counted as stutters:

> 'Behaviors such as rephrasing, repetition of phrases or whole-words, and pausing without tension are not counted as stuttering. Repetition of one-syllable words may be stuttering if the word sounds abnormal (shortened, prolonged, staccato, tense, etc.); however, when these single-syllable words are repeated but are otherwise spoken normally, they do not qualify as stuttering using the definition just stated.' (Riley, 1994: p. 4).

Second, SSI-IV provides flexibility in terms of how speech samples are obtained; they can be gathered from clinical settings, laboratory settings, or even from home. Moreover, the minimum sample length is specified. In the SSI-III, the minimum sample length is 200 syllables (Riley, 1994: p.9); however, this has been reduced to 150 syllables in the latest version (Riley, 2009: p.6). A study by Todd et al. (2014) investigated whether a 200-syllable sample is sufficient to obtain an SSI score, with the study revealing that the recommendation of a minimum 200-syllable sample for obtaining an SSI score is sufficient, and, in particular, that 200-syllable samples provide a more stable score when compared to shorter samples. However, the change to recommending a minimum sample of 150 syllables in Riley (2009) needs further evidence, and therefore it is recommended to adhere to the 200-syllable sample (Todd et al., 2014). Fourth, separate reader and non-reader assessment procedures are provided, which indicates the flexibility of the SSI in terms of the age ranges over which the assessment can be employed (Todd et al., 2014).

When evaluating the psychometric properties of SSI, there are some issues that need to be addressed. First, in terms of reliability, intra- and inter-rater reliability are the only forms of reliability estimation reported in the manual. Although the data presented in the manual indicates high levels of intra- and inter-rater agreement, the formulas used to calculate the percentage agreement (e.g. smaller/larger indices) are regarded as the weakest form of

reliability checking (Davidow & Scott, 2017). In this sense, such an approach to reporting reliability has been noted to be 'one of the least stringent measures of agreement' (Cordes & Ingham, 1994). More stringent statistics, such as other formulas of percentage agreement (e.g. percentage exact agreement or 1-point agreement) or methods such as limits of agreement (LOA) or standard error (SE) have not been incorporated in the manual. In addition, percentage agreement is a form of absolute reliability, and, to assess reliability comprehensively, it is useful to combine relative (e.g. the use of intraclass correlation coefficients [ICC]) and absolute reliability indices. With the exception of one recent study, no reports of the relative reliability of the SSI have been reported. In this regard, Tahmasebi et al. (2018) investigated the absolute and relative reliability of the Persian SSI-IV, using more stringent statistics, namely the ICC and SE. Ten judges rated 35 recordings of AWS and found satisfactory levels with regard to the relative reliability demonstrated by ICC, although less satisfactory levels of reliability were accomplished with absolute reliability indices.

The SSI-IV manual reports that frequency and duration values are expressed in stanine intervals, and these intervals are calculated from raw data. Thus, the raw data of frequency and duration 'within a given stanine have the same scale score value. This procedure reduces the effect of small differences on the test score' (Riley, 2009: p.15). Stanines typically represent a 9-point standard scale, with units (or values) that fall close to the mean representing small differences from the mean, whereas units (or values) that fall far from the mean represent greater differences (Davidow & Scott, 2017). However, no information is provided on how the stanine intervals were derived in a psychometric sense (Davidow & Scott, 2017). Moreover, the SSI-IV stanines (labelled as task scores for frequency measures and as a scale score for the duration measure) represent differing numbers of intervals (i.e. seven, eight, or nine) depending on which subscale is being used (Davidow & Scott, 2017). In addition, there is no information on the discriminative power of the Likert scale value that is summed to provide the physical concomitants score. It has been pointed out that the weakness of SSI lies in its low content validity, unsatisfactory inter-rater reliability, and the limited number of parameters used to assess stuttering severity (Healy, 1991).

Several studies have emerged dedicated to investigating the reliability of the SSI. In this sense, Hall et al. (1987) asked ten graduate students to rate two video recordings of two adults who stutter. Intra-judge reliability of the total score was satisfactory, as exhibited by correlation values of .816 and .889 for the two samples. However, inter-judge reliability was

less satisfactory. For the final severity rating, 70 and 80% inter-rater agreement was achieved for the two samples. In a second investigation by Lewis (1995), ten graduate students rated four samples of adults who stutter. The percentage agreement calculations stated in the SSI manual were used for analysis of inter- and intra-judge reliability, and the results produced were comparable to those stated in the SSI manual for sub-scores (i.e. frequency, duration, and physical concomitants), in addition to the total score. However, when comparing the ratings of raw counts of sub-scores, less satisfactory results were produced. Lewis also questioned the usefulness of the final severity rating (i.e. very mild, mild, and very severe), as it is calculated after several conversions, which may mask meaningful information, progress, or even regression in therapy. In addition, Lewis (1995) reported that a global rating of severity using a 5-point severity rating scale of samples corresponded with the final severity rating with of the SSI-III. In addition, Bakhtiar et al. (2010) investigated the reliability of the Persian version of SSI-III, wherein three judges scored 12 samples and intra- and inter-rater reliability were derived from the percentage agreement formula stated in the manual. Agreement was greater than 80% for frequency and duration, whereas low intra-rater (62.2%) and inter-rater (54%) reliability was found for physical concomitants. In this regard, the reliability levels were satisfactory, but the physical concomitants subsection must be interpreted with caution.

In a recent report investigating the reliability of SSI-IV, 12 judges rated four samples and data was analysed for inter- and intra-rater reliability of the SSI-4 sub-score, total score, and final severity rating (Davidow & Scott, 2017). As in Lewis (1995), the percentage agreement formula stated in the SSI-4 manual was incorporated to offer comparison, in addition to calculating exact agreement, within 1-point agreement, and within 2-points agreement. Intra- and inter-rater reliability across the sub-scores and total score were comparable to those reported in the manual. However, the new calculations of judge agreement (i.e. exact agreement, 1-point agreement, and within 2-points agreement) were different from the values in the manual for the three sub-scores, total score, and final severity rating (Davidow & Scott, 2017). Despite the contradictory evidence of the validity and reliability of the SSI, it remains a widely used measure, has been widely utilised to report details of people who stutter in more than 350 publications (Todd et al., 2014), and is regarded as the only measure of stuttering severity to use standardised procedures. In addition, the SSI has been translated into many other languages, including Persian (Bakhtiar, Seifpanahi, Ansari, Ghanadzade & Packman, 2010).

5.3 Relative and Absolute Reliability of the Arabic SSI-1V: An Investigation

5.3.1 Aim

The aim of the investigation is to evaluate the absolute (agreement) and relative reliability of an Arabic version of the SSI-IV.

5.3.2 Methods

5.3.2.1 Translation

For this study, the text of the SSI-IV examiner record form was translated into Arabic according to the Process of Translation and Adaptation of Instruments guidelines (WHO, 2015), whose steps include: 1) forward translation from English to the translated language and review of forward translation; 2) backward translation into English and review of the backward translation; 3) pre-testing the translated version of the test; and 4) final testing.

5.3.2.2 Forward Translation

Forward translation, from English into Arabic, was carried out by two experienced translators who were independent of the study. A panel consisting of three bilingual speech therapists (the primary investigator and two other SLTs) reviewed the SSI-IV record form items. Problems related to wording or clarity were considered and edited on the basis of their feedback after consensus regarding the translation. Amendments mainly related to the physical concomitants subsection, e.g. ‘noisy breathing’ was changed to ‘audible breathing’ and ‘foot tapping or swinging’ was changed to ‘tapping or rocking the foot’.

5.3.2.3 Backward Translation

Following forward translation, backward translation encompassed translating the SSI-IV back into English, which was conducted by a third experienced bilingual translator who was independent of this study. The translator’s role was to perform the backward translation and pinpoint discrepancies between the original test items and the Arabic translation. Discussion about discrepancies took place between the translator and the primary study investigator. No problems were reported at this stage.

5.3.2.4 Pre-Testing

The Arabic SSI-IV was distributed to three SLTs to trial. In this regard, the SLTs received a training session, with the aim of familiarising them with the translated version of the SSI-IV and explaining how to measure stuttering severity according to the SSI-IV manual. The SLTs

trialled the SSI-IV to school-age children who stutter in their own caseload for two weeks, and no problems with the final version of the SSI-IV were reported.

Finally, after receiving the test developer's agreement on translation accuracy, The Arabic version of the SSI-IV was used for the study (see Appendix 3). It should be noted that the SSI-IV norms are not validated for Arabic; however, it is thought that the use of US-norms is acceptable, as the SSI-IV is a non-language bounded test (Cook et al., 2011).

5.3.2.5 Participants

Eight school-age children who stutter were chosen at random from the main study (see Chapter 7).

Judges

The judges were four SLTs, including the main researcher, all of whom had clinical experience in different domains of speech and language therapy, including the assessment and treatment of children who stutter, with years of experience being 1, 2, 8, and 12. Although the judges selected had different levels of experience, this is thought to make findings more generalizable, as it reflects the community of SLTs (Karimi et al., 2014a). The SLTs received a training session, with an aim of familiarising them with the translated version of the SSI-IV and explaining how to measure stuttering severity according to the SSI-IV manual.

5.3.2.6 Procedure

Recordings of eight face-to-face sessions were randomly chosen from the main study, comprising approximately 25% of the data, which is regarded as sufficient to obtain reliability estimates. Although this sample size is small, it is comparable to previous published studies examining the reliability of SSI (Davidow & Scott, 2017; Lewis, 1995). Each participant had two speech samples; in case the child was a reader, one spontaneous sample and one reading sample was elicited, and, in case the child was a non-reader, two spontaneous samples in two situations were elicited (one sample was a conversation with the clinician and the other was a conversation with the parent). Topics for these spontaneous samples were those familiar to the children, e.g. hobbies, school, favourite TV shows, and holidays. In addition, picture plates were presented to the children to help elicit conversation. In total, 15 samples were elicited, where 1 non-reader participants recorded one sample only. The length of recordings varied depending on the stimulability of the child. The number of total syllables were above 200, thus following the recommended guidelines (Riley, 2009). More specifically, samples

ranged from 272 to 1,067 syllables, with a mean of 482. Judges incorporated the Arabic version of the SSI-IV to score the video recordings. Each sample was assessed twice by the judges, with a three-week interval in-between each assessment. To ensure consistency among scoring procedures, raters were only allowed to view recordings three times. In addition, the judges, with the exception of the main researcher, were blinded to the purpose of the study.

5.3.3 Data analysis

5.3.3.1 Intra- and Inter-Rater Reliability

As mentioned previously, the four judges assessed the eight recordings twice, with a three-week interval in between. The data from the first occasion was gathered to assess the inter-rater reliability of the eight recordings by the four judges. To assess the intra-rater reliability, recordings of the eight participants were re-assessed by the four judges three weeks later, and the scores of the three SSI-IV subcomponents, as well as the total score, were compared for the two assessment occasions. Recordings were anonymised by labelling them, and judges were not informed by the researcher that the samples were the same ones they had assessed on the first occasion.

Two forms of reliability were calculated for this investigation: relative reliability and absolute reliability (referred to as agreement). Relative reliability refers to the ability of a measurement to rank order individuals under conditions of repeated measurement (Batterham & George, 2003), whereas absolute reliability measures the closeness of scores to one another and to a hypothetical true score (Karimi et al., 2014a). In other words, relative reliability gives us an indication of whether the SSI-IV is suitable to distinguish between groups or to rank order them, or whether the SSI-IV is suitable to be used to detect temporal changes in participants' groups during clinical trials, whereas absolute reliability provides us with an indication of whether the SSI-IV is suitable for detecting small changes within an individual across time (Karimi et al., 2014a). Thus, the combination of absolute and relative reliability indices ensures a comprehensive assessment of the SSI-IV.

Intra- and inter-rater relative reliability were calculated by a two-way random effects model of intraclass correlation coefficient (ICC2, 1). In this regard, for the interpretation of the ICC values, we used Munro's taxonomy (Domholdt, 2005), where a value of .50–.69 is a moderate correlation, .70–.89 is a high correlation, and 0.90–1.00 is a very high correlation. For intra-

and inter-rater absolute reliability, the percentage agreement of exact scores and within 1, 2, and 3 sub-score values were calculated. This was preferred over Riley's percentage agreement calculation, which is derived by dividing the smaller value obtained by raters by the larger value and multiplying the quotient by 100 (Riley, 2009). Although using the latter measure may permit direct comparison with the SSI-IV manual, it suffers from shortcomings such as possibly providing different scores even when the judges are the same distance apart in their scorings. For example, in one sample, the first judge gave a rating of two and the second a rating of four, which results in a percentage agreement score of 50% (2/4 x 100 = 50%). In another sample, the first judge had a score of eight while the other had a score of ten, which resulted in an 80% percentage agreement (8/10 x 100 = 80%). Thus, although the judges were two points apart in rating both samples, the percentage agreements yielded were different (Davidow & Scott, 2017). Another problem with Riley's formula is that reliability estimates cannot be calculated when one judge has a rating of zero; in such a case, the rating should be omitted, thus causing a loss of information and resulting in a less informative reliability estimate.

5.3.4 Results

5.3.4.1 Inter-Rater Reliability

The intraclass correlation test indicated moderate to very high levels of inter-rater relative reliability between the four judges on the three subcomponents and the total score of the SSI-IV (see Table 3). In terms of absolute reliability, percentage agreement of the exact sub-score, percentage of agreement within 1 sub-score value, percentage within 2 sub-score values, and percentage within 3 sub-score values were calculated for frequency, duration, physical concomitants, and total score, in addition to the final severity rating (very mild, mild, moderate, severe, and very severe) (see Table 4). Judge agreement for the frequency sub-score was 43.75% for exact scores and 68.6%, 91.6%, and 95.8% respectively for within 1, 2, and 3 sub-score values. For the duration sub-score, agreement was 52.08% for identical sub-score but was notably higher for within 2 sub-score values (93.75%). For physical concomitants, judge agreement was 27% for identical sub-score and 58.3%, 79.1%, and 91.6% respectively for within 1, 2, and 3 sub-score values. Finally, for total score, judge agreement was 12.5% for identical sub-score and only 31.25%, 50%, and 68.75% respectively for within 1, 2, and 3 sub-score values. Final severity reached 100% within 1 sub-score value.

Table 3 Inter-Rater Relative Reliability of the SSI-IV between the Four Judges

SSI-IV Subcomponent	**ICC (95% CI)**
Frequency	.91 (.78 - .98)
Duration	.76 (.43- .93)
Physical Concomitants	.69 (.30 -.90)
Total Score	.91 (.76-.97)

Table 4 Inter-Rater Absolute Reliability of the SSI-IV between the Four Judges

SSI-IV Subcomponent	**PEA***	**Within 1 Sub-Score Value**	**Within 2 Sub-Score Values**	**Within 3 Sub-Score Values**
Frequency	43.75%	68.6%	91.6%	95.8%
Duration	52.08%	-	93.75%	-
Physical Concomitants	27%	58.3%	79.1%	91.6%
Total Score	12.5%	31.25%	50%	68.75%
Final Severity Rating	62.50%	100%	-	-

*PEA=Percentage of exact scores

5.3.4.2 Intra-Rater Reliability

The intraclass correlation test indicated high to very high intra-rater relative reliability for all the four judges on the three subcomponents and the total score of the SSI-IV (see Table 5). Tables 6–9 illustrate the intra-rater reliability of the SSI-IV of the four judges from Time 1 to Time 2. When the data was examined for the exact agreement of sub-scores, agreement percentages ranged from 62.5–75% for frequency, 62.5–100% for duration, 0–50% for physical concomitants, and 0–37.5% for total scores. Agreement increased significantly when the 1, 2, or 3 sub-scale value agreements were applied.

Table 5 Intra-Rater Relative Reliability of the SSI-IV between the Four Judges

SSI-IV Subcomponent	**ICC (95%CI)**			
	Judge 1	**Judge 2**	**Judge 3**	**Judge 4**
Frequency	.95(.78-.99)	.95(.8-.99)	.98(.94-.99)	.95(.8-.99)
Duration	.90(.54-.98)	.93(.72-.98)	1.0	.81(.32-.96)
Physical Concomitants	.88(.44-.97)	.86(.32-.97)	.94(.7-.98)	.94(.74-.98)
Total Score	.90(.6-.98)	.97(.89-.99)	.97(.87-.99)	.97(.89-.99)

Table 6 Intra-Rater Absolute Reliability of the SSI-IV of Judge 1

SSI-IV Subcomponent	**Percentage Agreement**			
	PEA*	**Within 1 Sub-Score Value**	**Within 2 Sub-Score Values**	**Within 3 Sub-Score Values**
Frequency	62.5%	75%	100%	-
Duration	62.5%	-	100%	-
Physical Concomitants	50%	75%	87.5%	87.5%
Total Score	37.5%	50%	87.5%	87.5%
Final Severity Rating	75%	100%	-	-

*PEA=Percentage of exact scores

Table 7 Intra-Rater Absolute Reliability of the SSI-IV of Judge 2

SSI-IV Subcomponent	**Percentage Agreement**			
	PEA*	**Within 1 Sub-Score Value**	**Within 2 Sub-Score Values**	**Within 3 Sub-Score Values**
Frequency	62.5%	87.5%	100%	-
Duration	87.5%	-	100%	-
Physical Concomitants	0%	100%	-	-
Total Score	0%	100%	-	-
Final Severity Rating	87.5%	100%	-	-

*PEA=Percentage of exact scores

Table 8 Intra-Rater Absolute Reliability of the SSI-IV of Judge 3

SSI-IV Subcomponent	**Percentage Agreement**			
	PEA*	**Within 1 Sub-Score Value**	**Within 2 Sub-Score Values**	**Within 3 Sub-Score Values**
Frequency	75%	100%	-	-
Duration	100%	-	-	-
Physical Concomitants	37.5%	87.5%	100%	-
Total Score	12.5%	62.5%	87.5%	100%
Final Severity Rating	100%	-	-	-

*PEA=Percentage of exact scores

Table 9 Intra-Rater Absolute Reliability of the SSI-IV of Judge 4

SSI-IV Subcomponent	**Percentage Agreement**			
	PEA	**Within 1 Sub-Score Value**	**Within 2 Sub-Score Values**	**Within 3 Sub-Score Values**
Frequency	75%	87.5%	100%	-
Duration	62.5%	-	100%	
Physical Concomitants	37.5%	75%	87.5%	100%
Total Score	12.5%	62.5%	100%	-
Final Severity Rating	100%			

*PEA=Percentage of exact scores

5.3.5 Discussion

The purpose of this study was to investigate the reliability of an Arabic version of the SSI-IV. Two levels of reliability were sought: relative and absolute reliability for intra- and inter-rater ratings.

In the current study, absolute inter-rater reliability, as calculated by percentage agreement, was not satisfactory in the case of percentage of exact agreement, nor within 1 or 2 sub-score values of the SSI subcomponents and the total score. More satisfactory levels of agreement were accomplished within 3 points of difference. A poor absolute inter-rater reliability is not unique to this study, as such results are in line with those of Davidow and Scott (2017), in which 12 judges measured seven speech samples of adults using the SSI-IV.

For the frequency sub-score, Davidow and Scott (2017) reported a percentage agreement of 43.75% for exact scores and 68.75% and 83.33% respectively for within 1 and 2 sub-score values. Our results were remarkably similar, with 43.75% for exact scores, and 68.6% and 91.6% respectively for within 1 and 2 sub-score values. For the duration sub-score, level of agreement in Davidow and Scott's (2017) study was 60.42% for exact sub-scores, which increased to 95.84% for within 2 sub-score values. Our results followed a similar trend, where

52.08% exact agreement was achieved and 93.75% agreement for within 2 sub-score values. For physical concomitants, Davidow and Scott (2017) reported an agreement of 27.08% for exact sub-scores and only 43.75% and 56.25% respectively for sub-scores within 1 and 2 values. Our study reported higher values, especially for within 1 and 2 sub-scale scores, where a 27% level of agreement for identical scores was reported, increasing to 58.3% and 79.1% for 1 and 2 sub-score values respectively. Finally, for total score values, Davidow and Scott (2017) reported a 27.08% percentage agreement for identical scores, a 39.58% agreement for within 1 sub-score, and 50% for within 2 sub-score values. Our figures were lower for identical and within 1 sub-score value, with 12.5% and 31.25% respectively. For the 2 sub-score value, we obtained the same percentage of 50%. As in Davidow and Scott's (2017) study, it was noted that, despite the poorer reliability for the sub-scores of the SSI-IV, improved reliability for the final severity rating was observed (see Table 2). Lewis (1995) has pointed out that the final severity rating may mask the individual subcomponents (frequency, duration, and physical concomitants); therefore, when relying on the final severity rating to track treatment changes, reliability tends to substantially improve, which is thought to cause a loss of information (Lewis, 1995).

When evaluating each subcomponent individually, the lowest absolute inter-rater reliability value for judges was the physical concomitants subcomponent, which is in agreement with previous investigations evaluating the reliability of the SSI (e.g. Bakhtiar et al., 2010; Davidow & Scott, 2017; Hall et al., 1987; Lewis, 1995). A possible explanation for the low levels of agreement found in this study and previous studies is the multi-tasking nature of scoring. Although the physical concomitants measure is scored after viewing the speech sample, it still entails the identification of sounds and facial and body movements, coupled with the counting of stutters and syllables and measuring duration all at once (Davidow & Scott, 2017). This simultaneous evaluation of multiple behaviours may offer an explanation as to why this subcomponent usually presents low levels of agreement (Davidow & Scott, 2017; Hall, 1987). In addition to the influence of multitasking, the discrepancy may be due to its subjective nature, as clear distinctions between each level of distraction are not provided (Riley, 2009; Tichenor, 2010). In other words, the way we differentiate a moment of stuttering with a distraction score of '1-not noticeable unless looking for it' and '2-barely noticeable to casual observer' is not clear. The same applies to a score of '3-distracting' or '4-very distracting', for what is distracting to one observer may not be as distracting to another (Tichenor, 2010).

On the other hand, analysis of ICC showed acceptable inter-rater relative reliability for the SSI-IV total score and its subcomponents. Poor absolute inter-judge reliability and acceptable relative inter-rater reliability have been previously reported for the SSI-IV specifically (Tahmasebi et al., 2018) and for other stuttering measures, such as the %SS and SR (Karimi et al., 2014a).

As for intra-rater reliability data, absolute intra-rater reliability scores were more satisfactory when compared to absolute inter-rater reliability. This may indicate that, when the SSI-IV is used by the same SLT, it is more reliable to measure minor changes in stuttering severity of a single client across time than when it used by different SLTs (Atkinson & Nevill, 1998; Karimi et al., 2014a).

This finding is of importance, especially in a clinical context, as, generally, the inter-rater reliability of overt speech behaviours of stuttering is important, given that the quantification of such behaviours is integral in both clinical and experimental research (Tahmasebi et al., 2018), e.g. when two different judges assess participants stuttering severity pre- and post-treatment in clinical trials. However, in clinical settings, intra-rater reliability may be more important, as stuttering severity ratings need to be internally consistent from session to session and across time (Tahmasebi et al., 2018).

As stated earlier, the acceptable levels of absolute intra-rater reliability for all judges in this study may indicate that the SSI-IV can be reliably used by a single clinician to assess minor changes in stuttering severity of a client across time. On the other hand, the poor absolute inter-rater reliability figures suggest otherwise, as, when interpreting the results of two different SLTs assessing stuttering severity of the same client across time, caution should be taken.

Satisfactory levels of relative reliability were achieved for both the inter- and intra-rater data, which may suggest that the SSI-IV could be used to distinguish between different participants, in addition to showing that the judges ranked the participants in almost the same order (Batterham & George, 2003).

In sum, the results of this study suggest that the Arabic version of the SSI-IV can be used by SLTs for research and clinical purposes. However, the compromised absolute inter-rater reliability suggests that the SSI-IV may not be appropriate for identifying minor changes in

stuttering severity within an individual across time, when different SLTs measure the stuttering severity of the same client.

6 Telepractice

6.1 What is Telepractice?

According to the American Speech-Language-Hearing Association (ASHA), telepractice is 'the application of telecommunications technology to deliver professional services at a distance' (American Speech-Language-Hearing Association, 2013). More specifically, it is a means of sharing health information and services using interactive video, audio, computer, and advanced telecommunication technologies to counsel, educate, assess, or treat clients in hospitals, schools, childcare centres, residential/non-residential healthcare facilities, and clients' homes (American Speech-Language-Hearing Association, 2013; Xue & Lower, 2010). While several terms exist, such as telehealth, telespeech, and telerehabilitation, the term telepractice is adopted in this thesis.

6.2 Telepractice in Saudi Arabia

In 1993, the Saudi government, by means of a royal decree, launched an e-Health Centre at King Faisal Specialist Hospital and Research Centre, which is one of the largest specialist hospitals in the Middle East, providing free medical care and accommodation for patients in need of treatment inside and outside Saudi Arabia (Dawoud et al., 2017). Telepractice services in Saudi Arabia permit healthcare providers to connect directly with leading specialists around the world. The centre delivers TP services, such as medical consultations and continuing medical education through international video conferencing and fibreoptic networks. Initially, the signed agreements for TP services only included a few hospitals in Saudi Arabia; however, years later, the number of connected sites for the national TP network has expanded to include twenty sites around Saudi Arabia, from major cities and vital rural areas, to advance telemedicine services and infrastructure and facilitate international connectivity for remote diagnostics and actual operations, in addition to voice and video conferencing services (Alyemeni, 2014; Dawoud et al., 2017; Jaber et al., 2014). In this regard, the use of TP services in King Faisal Specialist Hospital, along with other military and private hospitals, has proven successful (Jaber et al., 2014).

In 2018, Saudi Arabia announced its largest financial budget, which will have a major role in supporting the country's plans for expansion and diversification as part of its Vision 2030 programme, in accordance with which, the MOH expanded its telemedicine services by launching the 'Seha' application, which aims to cover all areas of Saudi Arabia (Saudi Arabia

Ministry of Health, 2018) in providing visual medical consultations and allowing clients anywhere to have face-to-face conversations with doctors across Saudi Arabia. The MOH is seeking to increase the number of clients using the application in order to bolster access to visual medical consultations on smartphones, in addition to undertaking efforts to employ cutting-edge technologies to enhance effective communication with service-beneficiaries, allowing them to get needed specialised medical consultations (Saudi Arabia Ministry of Health, 2018). The application is designed to enable audio-video communication from 8 am to 12 am during working days, as well as on weekends from 4 pm to 12 am. Accordingly, users can login to the application, communicate directly with a specialist, and have their cases diagnosed remotely. In this sense, specialists answer users' questions, providing needed medical consultations and facilitating necessary procedures (Saudi Arabia Ministry of Health, 2018). In addition, the MOH employs quality service programmes to maximise performance and to ensure the satisfaction of clients and health providers. In this regard, the MOH has also recently implemented the (e-health) initiative, which aims to enhance healthcare effectively and efficiently through the utilisation of information technology and digital transformation. Under this new initiative, the system supports health professionals, physicians, and nurses by providing them with patient information anytime and anywhere, including all documented health data. Users will be able to communicate with specialists and remotely receive the needed medical consultation. This new application is expected to reduce medical and diagnostic errors, as well as their incurred side effects, in addition to also enhancing online continuing medical education (Saudi Arabia Ministry of Health, 2018). Such investments can certainly be considered an endorsement of the validity, effectiveness, and future of telepractice within the Saudi healthcare system.

6.3 Internet Access

In order to incorporate telepractice services for children and their families, having access to the internet is imperative. At an international level, internet access is increasing, as one-third of the world's population is reportedly using it (Miniwatts Marketing Group, 2013). The internet was first made available to the Saudi public in 1999. Despite its relatively late introduction in Saudi Arabia, it has grown significantly in terms of connectivity (Alshahrani, 2016). This is evident in the dramatic increase in numbers of internet subscriptions: in 1999, there were 100,000 subscribers, which increased to 1 million by 2001 (Alshahrani, 2016). In 2013, the number of internet users further increased to 16.6 million, representing 55.1% of

Saudi Arabia's total population (Communications and Information Technology Commission, 2014; Ministry of Communications and Information Technology, 2003). In addition, it has been reported that nearly half of active internet users are below the age of 15 (Ministry of Planning, 2000). In addition, it is expected that, due to the growing availability of fibreoptic networks offering very high speeds, the demand for internet services will only increase (Communication and Information Technology Commission, 2012). Demand for broadband services has increased significantly compared to previous years, due to society's need for high bandwidth services, especially with the government's support for projects that require an efficient digital infrastructure (Communication and Information Technology Commission, 2012). In addition, many government services are now being provided through e-government transactions (Communication and Information Technology Commission, 2012).

6.4 The History of Telepractice

The introduction of telepractice can be traced to the early 1900s. In 1910, doctors used analogue telephone networks to transmit electrocardiograms and electro-cephalograms (Stanberry, 2000). By 1920, the provision of medical advice to people at sea via Morse code and voice radio was established (Stanberry, 2000). In the 1960s, two-way closed-circuit television systems were being used for the transmission of both images and consultations between health providers and patients, a method that is regarded as the precursor of video conferencing. Telepractice continues to be a major player in today's world – it is used in providing health services to rural populations and developing countries, as well as in clinics to transmit medical files, information, or images and to allow communication between healthcare providers (Stanberry, 2000).

6.5 The Scope of Telepractice

Telepractice is currently used for a wide range of health issues (DePalma, 2009). It has been utilised in the assessment and treatment of children and adults in different fields of medicine and rehabilitation, including psychology, speech and language therapy, and audiology (Houston, 2013). In an international review of telepractice, encompassing 578 reports published between 1990 and 2003, it was stated that, due to increased demands for accessible and efficient healthcare, an aging population, and the obstacles in retaining home-visiting health professionals, the development of TP services took place in developed countries. Low-cost technologies and high internet availability have made telepractice a more affordable treatment choice. However, limitations have also been reported, including the lack of

protocols, guidelines, and standards; the need for evaluation; ethical considerations; usability; and economic considerations (Koch, 2006).

6.6 Telepractice in the Context of Speech and Language Therapy

Telepractice is used in many countries around the world, including the United States of America, Canada, Greece, Ireland, the UK, and Japan, for a range of impairments, from neurogenic communication disorders and paediatric speech and language impairments to dysphagia and stuttering (Mashima & Doarn, 2008). As a result, many speech therapy professional organisations have published position statements regarding telepractice (e.g. American Speech Language Hearing Association, 2013; Canadian Association of Speech Language Pathologists and Audiologists, 2006).

In Saudi Arabia, there are no existing laws or specific regulations regarding TP services, and thus no position statements or guidelines exist. Therefore, it is assumed that the same regulations and laws that exist in traditional health services generally pertain to telepractice. In other words, SLTs who utilise TP must conform to the same regulations and laws as those who practice face-to-face.

6.7 Telepractice Delivery Models

As research has progressed in the field of telepractice speech-language therapy intervention, three different-yet-related models of TP speech and language therapy have emerged: synchronous, asynchronous, and hybrid methods of TP intervention (Mashima & Doarn, 2008).

6.7.1 Synchronous

In this service delivery mode of telepractice, the SLT and the client can interact in a direct manner (real-time) via audio or video technologies, such as the telephone or video conferencing, for the purpose of diagnosis, assessment, or intervention. Common modes for connection include audio conferencing, chat, white boarding, instant messaging, and application sharing. One example of synchronous services includes the real-time conduction and interpretation of video- fluoroscopic studies and assessment and treatment via video conferencing.

Video conferencing can be conducted in several ways, given that broadband connections have become increasingly widespread and available to the public (Dudding, 2009). Video conferencing technology includes a range of options, from dedicated video conferencing

equipment, such as products from companies like Polycom, to commercially available equipment such as laptops, tablets with built-in webcams, microphones, and speakers (Flemings, Brown & Houston, 2013). Once the hardware is obtained, the only additional requirements are an internet connection and video conferencing software (e.g. Skype and FaceTime) (Fischer et al., 2016).

Communication, problem solving, and behaviour modification are of a dynamic nature, and, therefore, in the field of SLT, synchronous transmission for 'real-time' interactive evaluation and treatment of communication disorders is often desired (Mashima & Doarn, 2008). The focus of this thesis is on video conferencing, wherein visual and auditory information is transferred real-time via the internet, enabling the SLT to directly conduct activities with a client with the provision of immediate feedback, regardless of how physically far apart the client and the SLT are (Theodoros, 2008).

6.7.2 Asynchronous

The asynchronous, or store-and-forward, TP delivery model does not require the clinician and client to communicate directly. Thus, in this mode, samples are collected at a convenient off-site location. Samples can include video recordings, audio recordings, or other artefacts, such as written documents (e.g. medical records), test protocols, client data (e.g. test protocol results and treatment performance), and still pictures (ASHA, 2013). Once this data is collected, it can be sent to the clinician through a variety of carriers – the internet, telephone modems, or fax machines (ASHA, 2013).

6.7.3 Hybrid

Hybrid TP applications include combinations of synchronous, asynchronous, and/or traditional face-to-face services (ASHA, 2013). The SLT plays a vital role in deciding what model is used, with the aim of maximising clients' outcomes, bearing in mind cost, ease of use, and overall patient satisfaction (Mashima & Doarn, 2008; Xu & Lower, 2010).

6.8 Past Reviews of Telepractice Applications in the Field of Speech and Language Therapy

Hill and Theodoros (2002) conducted a comprehensive review of the literature of TP applications in speech and language therapy. Thirteen empirical studies were included, of which two related to stuttering therapy, while the others reported on the application of TP services to childhood speech and language disorders, acquired neurological communication

disorders, and voice disorders. The authors reported that the results of these studies were encouraging, and most of the studies included in the review concerned the overall efficacy of telepractice. Up to the time of their review, it was noted that most TP applications of SLT had involved investigating new methods with 'proof of concept' rather than substantial empirical studies. The authors concluded that these studies lacked sufficient details to facilitate replication and encompassed small sample sizes. They also indicated that no economic analysis of the applications had been conducted and that SLT outcome measures were lacking. It was argued that evaluation of the technology used is important, but no studies were identified during the time the review was being published. Finally, they noted that TP interventions were not applied to limited speech and language disorders. It was suggested that future studies should investigate client and SLT satisfaction, evaluate processes and procedures, and include a cost-benefit analysis as part of the assessment of client outcomes. It was also pointed out that there is a need to develop protocols and guidelines in different environments for specific client groups, with consideration of the types of technology available to service providers and the wider community, including clients.

Mashima and Doarn (2008) conducted an extensive literature review on the application of telepractice in SLT with adults and a small number of studies with children. They reviewed 40 studies investigating disorders relating to adult neurogenic communication, fluency, voice, dysphagia, and childhood speech and language disorders. The authors focused their discussion on several topics, such as the technology used in TP delivery, client and SLT satisfaction, advantages of telepractice, challenges and barriers to the application of TP, and future directions. It was suggested that telepractice is a feasible and effective method for providing remote SLT services. However, it was pointed out that the studies reviewed were primarily comprised pilot studies and anecdotal accounts of TP applications rather than large, well-controlled, randomised clinical trials.

Reynolds, Vick and Haak (2009) conducted a narrative review of 29 studies, which were analysed using a quality assessment checklist. These studies focused on assessment and treatment with both adult and child populations as well as an unspecified population. The authors concluded that the results achieved through telepractice and in-person service delivery models were equivalent, though many of the studies commented that telepractice cannot completely replace face-to-face services; rather, it is recommended as a complement to face-to-face services. These findings were in line with the Theodoros' (2012) review, which

encompassed 19 studies regarding adult neurogenic communication, voice, stuttering, dysphagia, and laryngectomy disorders and four studies about paediatric speech, language, and literacy disorders.

It should be noted that the aforementioned studies were primarily focused on adults, with less emphasis on the paediatric population. In response to this bias, a systematic review was conducted by Wales, Skinner and Hayman (2017), with the aim of investigating if telepractice SLT interventions are as effective as traditional face-to-face delivery for primary school-age children with speech and/or language difficulties. Seven articles were included in the review, and results revealed that both TP and face-to-face participants made significant and similar improvements when treatment effects were assessed through five of the six outcome measures. Findings showed that there is limited but promising evidence to support telepractice for delivering speech-language pathology intervention services to school-age children. However, the authors commented that more rigorous study designs are required to support the efficacy of telepractice for this population.

Finally, a review by Taylor et al. (2014) focused on the role of telepractice in providing assessments for the paediatric population. In their systematic review, five articles were chosen for analysis. These studies aimed to assess the efficacy of TP using method comparison techniques and thus to investigate the validity and/or reliability of speech and language assessment via telepractice. One study considered the screening of speech and language using formal tools (Ciccia et al., 2011), while the remainder conducted full assessments using both formal and informal methods (Eriks-Brophy et al, 2008; Waite et al., 2006; Waite et al., 2010; Waite et al., 2012). Four papers investigated the assessment or screening of speech, particularly articulation, oromotor function, and speech intelligibility (Ciccia et al., 2011; Eriks-Brophy et al., 2008; Waite et al., 2006; Waite et al., 2012), and three papers examined assessment or screening of receptive and expressive language (Ciccia et al., 2011; Eriks-Brophy et al., 2008; Waite et al., 2010).

The authors noted that comparing studies was not an easy task, as they varied in terms of participant groups, assessment tools used, statistical analysis, and the TP equipment incorporated. They criticised the statistical analysis utilised to measure agreement in the identified studies, as they incorporated the use of Pearson's *r*, which is indicative of correlation not agreement, percentage agreement, which is regarded as an informal measure, or kappa statistics, which can be difficult to interpret, with only two studies using more

rigorous statistics to calculate agreement, particularly the Bland-Altman Limits of Agreement method (Bland & Altman, 1999).

The studies reviewed indicated some evidence that telepractice can be used to make valid assessments of oromotor function, speech intelligibility, and language. It was found that articulation screening via TP is valid; however, results were contradictory in the case of full articulation assessment. As for reliability, intra- and inter-rater reliability was satisfactory for all speech and language tasks, except for oromotor tasks; however, oromotor tasks had reduced reliability, even when assessed face-to-face. There were contradictory findings between the two modalities for individual oromotor tasks, judgment of individual speech sounds, detection of pluralisation, and discriminating between similar sounding words, which are regarded as clinically important (Taylor et al., 2014).

Evidence regarding satisfaction was extremely limited, with only one study reporting parental satisfaction, and no studies reporting child or clinician satisfaction at the time of writing. The authors concluded that, if evidence exists to support the validity and reliability of speech and language assessment via TP, it is probably not yet adequate to guide clinical practice or policy development (Taylor et al., 2014).

It was noted that, among these reviews, no research has been published on the viability of video conferencing for the assessment of children who stutter, although there have been studies regarding stuttering therapy conducted via TP.

6.9 Potential Benefits and Limitations of Telepractice

6.9.1 Potential benefits

6.9.1.1 Benefit of meeting unmet needs

As stated earlier, shortages of SLTs in Saudi Arabia continue to be a problem, an issue shared by many countries around the globe (Wylie et al., 2018). One of the foremost advantages of telepractice is that it improves access to services by eliminating the need for proximity. In the United States, medical rehabilitation facilities are faced with vacancies in speech-language therapy positions, resulting in individuals with communication disorders not receiving necessary intervention (Edwards, Stredler-Brown & Houston, 2012; Juenger, 2009; Polovoy, 2008). To tackle this issue, many school districts, hospitals, and private practitioners are using telepractice to provide SLT services for individuals who would otherwise go without these services (Brady, 2007; Juenger, 2009; Polovoy, 2008; Tucker, 2012).

Shortages of SLTs have significantly impacted upon school settings in rural communities (Crutchley & Campbell, 2010). Telepractice has been successfully used in certain school districts to provide students with SLT services (Grogan-Johnson et al., 2013). In a description of TP use in public schools, (Polovoy, 2008), institutions using basic internet technology to conduct SLT sessions via video conferencing, or those that sign contracts with private companies to do so, have reported successful outcomes for individuals who otherwise would have been forced to go without treatment.

In the case of Saudi Arabia, SLT services in schools are non-existent; the only way to access SLT services is either in healthcare settings or private practice. Since in-person treatment is not an option for schools, telepractice could fill the gap by connecting SLTs to schools, thus meeting the communication needs of students.

In addition to improved access and reducing time and distance concerns, telepractice may assist in the coordination and continuity of care as it permits the transferral of comprehensive information, regardless of physical location (Buchanan, 2002). It can also aid clients in receiving weekly services that otherwise would have been unviable due to lengthy travel times resulting from traffic congestion (Verdon et al., 2011). In many cases, clients prefer telepractice to the face-to-face modality due to its convenience, time savings, and reduction in transport costs (Mashima & Doarn, 2008; Theodoros, 2011). In addition, TP treatment can have stronger generalisation gains than face-to-face clinic-based treatment (Burgess et al., 1999; Mashima & Doarn, 2008; Theodoros, 2011), making the TP delivery model beneficial not only to those from rural and remote areas but also metropolitan clients who find it problematic to attend face-to-face sessions.

Other situations where telepractice may be preferable to the face-to-face delivery model is, for example, in the case of assessment, where changes in routine, such as travelling or interacting with strangers in a waiting room, may be stressful, especially for individuals with developmental disabilities (Tyler & Tolbert, 2002). Exhibiting such anxiety could be misleading, as said behaviour may not accurately reflect how a client thinks, acts, and responds on a day-to-day basis. The provision of services to clients in their functional environments, considered best practice in many areas of rehabilitation (McCue, Fairman & Pramuka, 2010), is supported by the World Health Organization (WHO) intervention framework (WHO, 2001). Thus, telepractice has the potential to improve client outcomes by

targeting their functional environment, sustaining services, facilitating self-management, and reducing costs.

6.9.1.2 Benefit of reduced cost

Telepractice technology offers access to, and the availability of, expertise at minimal costs (Miller et al., 2006). Despite the many studies incorporating video conferencing using custom-built systems, there is also the option of using low-cost consumer-grade equipment, e.g. personal computers or tablets with in-built cameras using home internet connections. However, there is limited evidence supporting the use of PC-based video conferencing for disorders that require high-fidelity audio signals, such as speech and voice, including stuttering (Keck & Doarn, 2014). It has also been noted that, although TP treatment may be less expensive than face-to-face SLT services, for some families the contrary is the case, particularly when there are limits on financial reimbursements for TP services by private health insurers.

6.9.2 Potential limitations

6.9.2.1 Reliability and validity

With regard to assessment, the validity and reliability of a new TP assessment is an important concept to address when evaluating services, as it is considered useless if the TP assessment does not evaluate the same construct as a face-to-face assessment, regardless of its potential benefits (Russell et al., 2017).

One factor that may threaten the reliability and validity of an assessment carried out via the TP model is the chance of interaction between the construct being tested and the medium utilised to test it (Buchanan, 2002). Therefore, the veracity of this model should be explored using rigorous designs prior to its implementation in clinical settings.

Another factor that may affect validity and reliability is how comfortable clients are with technology, as this could lead some to respond differently to measures when compared to face-to-face services (Sharma et al., 2013).

6.9.2.2 Security, privacy, and confidentiality

Privacy and confidentiality of clients is central when providing healthcare services, including SLT services (Cohn & Watzlaf, 2011). The provision of SLT services via the TP model poses

unique challenges in terms of privacy and confidentiality when compared to face-to-face service provision.

For example, with regard to assessment, SLTs produce written reports that are stored and used according to ethical guidelines. In the case of TP assessment, instead of a written report, an electronic document is produced along with a video recording of the session if needed, which is captured by the video conferencing software used to carry out the session (Houston, 2013). Digital versions of sessions may be helpful for interventions as well as for educational and research purposes. However, they pose confidentiality and privacy challenges.

In addition, during video conferencing, a client's privacy may be physically breached, for example, if non-clinical personnel can hear part of a session conducted via telepractice due to an open door, the transmission of sound between rooms, or a colleague unexpectedly entering the room, or as a result of the presence of a non-authorised person (e.g. a family member) hearing and/or seeing part of the sessions at the client's side (Watzalf & Cohn, 2011). Therefore, clients should report who will be present at their side when they receive TP services, a camera may be used to scan the environment of the SLT to confirm privacy, and the identification of persons in rooms at both sides should take place before conducting a telepractice session (ASHA, 2013).

Telepractice sessions are usually conducted via video conferencing software such as Skype and FaceTime, and the use of such software, although it may seem convenient and affordable, could contradict the Health Insurance Portability and Accountability Act of 1996 (HIPAA), as they pose security threats, given that not all video conferencing software protects data privacy or is securely encrypted. Therefore, prior to the use of any video conferencing software, SLTs should consider whether the video and or/audio transmissions and any other information shared via video conferencing will be kept private and secure and will comply with HIPAA requirements (Watzalf & Cohn, 2011).

As a risk management measure, SLTs should obtain informed consent from their clients, which fully describes the equipment and services to be delivered, the differences between services delivered via telepractice as opposed to services delivered face-to-face, the client's right to return to face-to-face services at any time, any adaptions that will be made in assessment protocols, and potential confidentiality and security issues (ASHA, 2013).

6.10 Acceptability of and Satisfaction with the Telepractice Model

As stated earlier, telepractice may improve client outcomes and help alleviate barriers of distance and travel. However, for this model to be effective, it is critical to evaluate end users' insights into the acceptability of and satisfaction with TP services in order to facilitate their ultimate clinical uptake and success (Krupinski et al., 2008).

In the case of assessment, which is the focus of this thesis, few studies obtained information from children and/or their parents regarding their satisfaction with TP assessment. In Ciccia et al. (2011) and Whitehead et al. (2012), high levels of satisfaction were reported. Parents expressed that they were willing and interested in using TP services in the future and were happy about the audio and video quality of the sessions. Sutherland et al. (2017) reported that parents and their children were comfortable with the TP assessment. Although most parents were positive regarding their experiences, a few parents expressed their concerns regarding telepractice and reported that they would not be interested in pursuing this service delivery model in the future. This finding is contradictory to that of Ciccia et al. (2011), where parents preferred to receive screening services via telepractice rather than attending a face-to-face appointment.

Sutherland et al. (2017) also surveyed SLTs regarding telepractice audio and image quality, in addition to issues around the child, clinician, and technological aspects. It was reported by SLTs that no child, SLT, or technological issues precluded the assessment. SLTs ratings of audio and video quality were good for 74% and 83% of the time respectively. SLTs reported that they overcame technological issues such as low volume and words cutting out by repeating instructions and speaking more loudly. However, it is thought that results from a certain study cannot be generalised to others, as each is influenced by the assessment procedure incorporated, technology chosen, internet connectivity, infrastructure, and human resources. These factors, coupled with cultural factors, may influence perceptions about and acceptance of the adoption of the TP model to provide SLT services (Mansouri-Rad et al., 2003; Williams, May & Esmail, 2001). In addition, to our knowledge, there has been no investigations of client or parent satisfaction with stuttering assessment services delivered via telepractice. Furthermore, the aforementioned studies only collected satisfaction information post-assessment, with no study examining both the pre- and post-assessment perceptions. Considering that pre-conceptions of a service may have an impact on the willingness of clients to take up the service, this is thought as being an essential additional aspect that should

be part of an evaluation of patient perception and satisfaction with TP services (Sharma et al., 2013).

6.11 Telepractice Applications in Speech and Language Therapy

Some studies investigated assessment procedures, while others focused on treatment. Some targeted adults, with fewer studies focusing on children (Houston, 2013).

Research studies have investigated the feasibility, efficacy, and effectiveness of using telepractice in a variety of settings; they have used it to provide screening, assessments, and treatments to several adult and child populations. In the following section, the results of these studies are discussed, with the final section dedicated to the use of telepractice in stuttering.

6.11.1 Telepractice applications of adult speech and language disorders

6.11.1.1 Assessment

A study by Brennan et al. (2004) aimed to measure the performance of adults with acquired brain disorders on standardised SLT assessments conducted under two conditions: face-to-face and video conference settings. For the purpose of the study, 40 participants were assessed using story-retelling tasks. The two stories were selected randomly, and all participants were randomised and tested in both conditions. Additionally, the authors surveyed participants to measure their satisfaction with the TP service provided. Results indicated no significant differences in outcomes and a high agreement (93%) between the two conditions across all subject variables (e.g. age, gender, and experience with technology). Additionally, participants expressed a high level of interest in using telepractice in the future. It was concluded that story-retelling performance was not affected by settings but that further research is warranted (Brennan et al., 2004).

Another study by Palsbo (2007) investigated the equivalence of video conferencing assessment of 24 post-stroke patients diagnosed with aphasia using the Boston Diagnostic Aphasia Examination (BDAE). A double-crossover agreement study was conducted, and participants were randomised to a TP or face-to-face administration. Results indicate that clients' functional communication could be assessed reliably using the TP model (95% agreement in each functional communication measure regardless of assessment site).

In another study by Hill et al. (2009b), which examined the influence of the severity of aphasia on TP assessment, the Boston Diagnostic Aphasia Examination (BDAE) and Boston Naming Test (BNT) were used for 32 patients with aphasia who were assessed

simultaneously. Participants were grouped according to their aphasia severity and were randomly assigned to face-to-face and TP conditions. In addition to the random allocation of participants, two SLTs were randomly assigned to one of the two assessment conditions. Results revealed that inter-rater agreement, in most cases, was above 90%, and it was concluded that the efficiency of telepractice assessment was not influenced by the severity of aphasia, except for a few tasks (such as the ability to assess naming and paraphasia) (Hill et al., 2009b).

The validity and reliability of the dysarthria assessment via telepractice was also explored. In this regard, 24 adults with dysarthria were assessed using standardised and informal assessments conducted via a purpose-built TP system. As in the aforementioned study, participants were assessed simultaneously via the TP system and face-to-face conditions. Results indicated clinically acceptable inter-rater agreement of 80 to 100% between SLTs in face-to-face and TP conditions, in addition to high intra- and inter-rater reliability of test components. It was concluded that assessment of dysarthria using telepractice is valid and reliable (Hill et al., 2009a).

The validity and reliability of assessing apraxia of speech using telepractice was investigated. Assessments administered face-to-face and via TP were scored simultaneously by two SLTs. The results of the study indicated no significant differences between the subtest scores for the two environments. While considering the small sample size, the authors suggested that reliable and valid assessment of apraxia was feasible using telepractice (Hill et al., 2009).

Ward et al. (2009) simultaneously assessed ten post-laryngectomy patients for swallowing, stoma, and communication status via a distant clinician using a purpose-built video conferencing system and via a second clinician who was present with the patient. A satisfaction questionnaire was also completed. Results revealed high agreement between the two SLTs and high satisfaction for both patients and clinicians. The authors emphasised that image quality obtained via the camera was rated lower than direct observation, but it was sufficient to assess the stoma and the status of the voice prosthesis.

Ward et al. (2011) investigated the validity of conducting dysphagia assessments with 40 participants. Participants were assessed simultaneously by a face-to-face SLT and a TP SLT via a custom-built video conferencing system. Assessment material, comprised Clinical Swallowing Examination protocol, was delivered via the video conferencing system, and an assistant at the patient's end was present to facilitate the assessment. Results revealed

acceptable clinical agreement in both face-to-face and TP conditions. While the authors supported the validity of conducting assessments through telepractice, they pointed out that complex diagnostic conditions should be evaluated in a traditional face-to-face setting.

6.11.1.2 Therapy

In a study by Lasker et al. (2010), a 28-year-old patient with aphasia and apraxia received SLT therapy sessions that lasted 50 minutes, four times per week, for a total of 12 months. Two sessions per week were conducted in person and two sessions via Skype. Results revealed that the outcomes were similar in both conditions (i.e. TP and face-to-face) (Lasker et al., 2010).

Another study by Goldberg, Haley and Jacks (2012) examined the effects and generalisation of modified script training intervention via telepractice. Two aphasic patients, trained in relevant scripts, participated in a three-day-per-week intervention, whereby training on each script lasted for three weeks. Participants enjoyed a combination of face-to-face and TP sessions delivered via Skype. Both participants demonstrated improvement, and it was concluded that telepractice is a viable method for script training (Goldberg, Haley and Jacks, 2012).

A study by Mashima et al. (2003) investigated the effectiveness of delivery of voice therapy via telepractice, compared to conventional face-to-face delivery. Seventy-two (34 male and 38 female) individuals took part in this study, were randomly assigned to either the TP setting or the face-to-face setting, and then matched according to their diagnostic category. For each group, baseline measures were obtained, and treatment methods and facilitating techniques were consistent for both face-to-face and telepractice (measures included the Perceptual Judgments of Voice Quality). Each participant recorded both pre- and post-treatment samples. Each individual's sample was then presented with its matched diagnostic pair from the other group, and two SLTs determined which sample demonstrated better voice quality. Live voice samples of a sustained /a/ were also obtained and analysed on the Visi-Pitch II system. No significant differences were found between the face-to-face and TP settings; therefore, it was concluded that telepractice would be an effective model to deliver voice therapy.

Thirty-four patients with Parkinson's disease were treated for voice disorders via telepractice services. The data for pre- and post-treatment measures revealed significant progress, with no

major difference between outcomes for participants assigned to face-to face and TP conditions (Constantinescu et al., 2011).

6.11.2 Telepractice applications to paediatric speech and language disorders

6.11.2.1 Assessment

Fairweather et al. (2004) investigated the assessment of the language, speech, and oral motor functioning of 13 children aged 6 to 14, using both informal and standardised tests via telepractice, with formal assessments including the Goldman-Fristoe Test of Articulation (Second Edition) (GFTA-2) and the Clinical Evaluation of Language Fundamentals (Third Edition) (CELF–3) and informal assessment encompassing an informal conversational speech sample analysis and oromotor assessment. Assessments were conducted and rated via video conferencing, in addition to being simultaneously rated by another SLT who was present with the child. The face-to-face SLT also had the role of adjusting equipment and stimulus material upon request from the TP clinician. Results were encouraging, as there was total agreement between the two SLTs 69% of the time on the severity of articulation disorders and 92% on the severity of language disorders. Further, 69% agreement was achieved for individual subsets of the CELF-3, with the authors pointing out that discrepancies were due to subjective scoring and the difficulty to accurately differentiate spoken words by children; it was often not clear whether the difficulty in discerning spoken words resulted from poor audio quality during video conferencing or the low intelligibility from the children. Regarding articulation assessment, it was noted that discrepancies in SLTs' judgment were evident in participants with severe articulation disorders and with certain classes of sounds, which made the authors recommended that the TP system should be used for mild cases as opposed to moderate or severe ones. Finally, with regard to informal assessment, particularly oromotor assessment, the authors pointed out that assessment was not smooth, as problems such as sufficient viewing angles and poor lighting made some tasks difficult to complete (Fairweather et al., 2004). This raises an obvious yet noteworthy point: not all telepractice assessment tasks or tools can be administered in an effective manner.

A pilot study assessing oral-motor structure and function and articulation was carried out by Waite et al. (2006). This study included six children between the ages of four and nine who had speech impairments. These children were simultaneously assessed using a video

conferencing system. The results were positive, in that the agreement between face-to-face and video conferencing was 91% for the oral-motor tasks, 100% for speech intelligibility, and 92% for single-word articulation. As in the aforementioned study, it was reported that discrepancies amongst ratings were probably due to technical and methodological problems, such as positioning the children to achieve camera angles (Waite et al., 2006). The authors argued that a movable web camera with a zoom function may increase reliability. Another reason for the discrepancy between scores may have been due to the subjective nature of oral-motor ratings rather than the TP model. The researchers stated that further investigation of other client groups and larger sample sizes was necessary (Waite et al., 2006).

Ciccia et al. (2011) conducted a feasibility study of providing speech, language, and hearing screening to 411 children up to the age of six using Skype at two primary care clinics in Cleveland, Ohio. The children were living in heavily populated urban areas and were screened during scheduled doctor visits. Hearing screenings included administering tympanometry for children of all ages, Distortion Product Otoacoustic Emissions (DPOAE) for children up to three years old, and behavioural audiometry for children aged three to six. Speech and language screening consisted of administering the Receptive-Expressive Emergent Language Test (Third Edition) (REEL-3) for children up to three years old and the Screening Kit of Language and Development (SKOLD) for children two-and-a-half to six years old. Results revealed high levels of family satisfaction with the screenings, and, out of the children who failed screenings, 72% of parents scheduled appointments for comprehensive evaluations. Good reliability for both hearing and speech-language screenings was reported; however, the reliability calculation was comprised small samples of participants, e.g. n=10 participants were included in the reliability analysis for speech and language screenings.

Eriks-Brophy et al. (2008) delivered speech and language assessments to seven aboriginal children aged between 4 and 12 in a remote area of Canada via telepractice. The authors used the GFTA-2 for speech assessment and the Peabody Picture Vocabulary Test (Third Edition) (PPVT-3), Preschool Language Scale (Fourth Edition) (PLS-4), Clinical Evaluation of Language Fundamentals (Fourth Edition) (CELF-4), and Expressive One-Word Picture Vocabulary Test (EOWPVT) for language assessment.

High mean percentage agreement between conditions for language assessment occurred (98–100%). However, a less favourable percentage agreement was achieved for GFTA-2, while

discrepancies were noted in CELF-4 (detection of plurals) and GFTA-2 (judgement of individual speech sounds).

The results of this study demonstrated the potential of telepractice for language assessments, but it was unclear what TP system was used, and the very small numbers meant that replication was required and generalisation was precluded. It was noted that there was an increased need for repetition during TP assessment, which consequently increased the time of the assessment, which, in turn, may threaten the validity of a standardised test, especially if that test allows a limited number of repetitions (Eriks-Brophy et al., 2008; Taylor et al., 2014).

The authors observed that children's participation improved due to their enthusiasm and curiosity about TP technology. This was apparent when children were asked about their preferred screening method, in contrast to their parents, who preferred face-to-face assessment. This discrepancy in opinions illustrates the importance of conducting research into understanding both parents' and children's satisfaction with telepractice, which is one of the focal points of this thesis; as the SLT profession is family-centred in nature, understanding both views will aid in balancing the service preferences of both parents and their children when planning assessment services (Crais, Roy & Free, 2006; Eriks-Brophy et al., 2008; Taylor et al., 2014).

A study by Waite et al. (2010) aimed to validate the use of CELF-4 via telepractice using a custom-built internet-based TP application with store-and-forward capabilities and touch screen at the child's end. Picture stimuli were in digital format, with the use of audio recording of the stimuli for one subtest assessment and live presentation of the remaining three subtest materials. Twenty-five children between five and nine years old were assessed using the core language subtests of the CELF-4. Each child was simultaneously assessed via telepractice and face-to-face, while assessments were administered by either a TP or face-to-face SLT but were simultaneously rated by both SLTs. Results were positive, as no significant difference was found between the telepractice and face-to-face total raw scores and scaled scores for each subtest of the CELF-4. Weighted kappas revealed very good agreement on the individual items, total raw scores, scaled scores, core language score, and severity level. As for intra- and inter-rater reliability by means of intraclass correlation analysis, very good reliability was achieved on all measures. However, Taylor et al. (2014) indicated that the study had limitations since it was conducted in two rooms in the same building of a

university clinic under ideal conditions, rather than in a real-world setting. In addition, the custom-built equipment and completion of the study in a research setting limited its applicability to clinical settings, while no client or clinician satisfaction was reported on. Finally, the studies did not include comparisons of child behaviour across face-to-face and online conditions (Sutherland et al., 2017).

A more recent application of CELF-4 via telepractice took place in a rural area of New South Wales, Australia (Sutherland et al., 2017). A TP SLT delivered the core subtests of the CLEF-4 to 23 school-age children aged between 8 and 12 in a remote hub using video conferencing. A second SLT, located in the same room as the student, facilitated the assessment (e.g. turning on the computer and logging into the site) and co-scored the subtests. Compared to Waite et al. (2010), this study went one step further by including a behaviour observation checklist for children, a satisfaction questionnaire for parents, and SLT ratings of audio and video quality in the TP assessments. Finally, the study used commercially available equipment and software rather than custom-built software. Good agreement on all measures of the CELF-4 was found, including the severity of language disorders. In addition, inter-rater reliability in the TP and face-to-face conditions was high (correlation coefficients ranged from r=0.96–1.0 across the subtests). Comparable levels of attention, distractibility, and anxiety were noted across the two conditions (TP vs. face-to-face), measured by the behaviour checklist, and parent reports were largely positive. It should also be noted that, compared to previous studies, the agreement analysis incorporated more rigorous statistics, as they utilised the Bland and Altman method (Bland & Altman, 1999).

6.11.2.2 Therapy

Grogan-Johnson et al. (2010) conducted a pilot study in which they provided SLT services incorporating telepractice to four schools in rural Ohio and compared the progress made by school-age children under two conditions (face-to-face and telepractice). The children, divided into two groups, either exhibited articulation, language, and/or fluency disorders. In the first group, 17 children received treatment via telepractice for four months and then subsequently received face-to-face therapy for four months, whereas, in the second group, 17 children received face-to-face treatment for four months and then subsequently received TP treatment for four months. The outcome measures were student progress, participant satisfaction, and any interruptions to service delivery. Results, indicated by means of quarterly progress reports, revealed that students from both groups made similar progress, and

there was no significant difference in GFTA-2 score between students in the two treatment groups. In addition, satisfaction surveys indicated that students and parents supported the use of telepractice to deliver SLT services.

In a follow-up study by Grogan-Johnson et al. (2011), telepractice delivery of speech sound intervention for 13 school-age children between the ages of 6 and 11 was compared with standard face-to-face delivery. Students received speech sound intervention using computer-based materials provided either through TP or face-to-face intervention. Students in both service delivery models made significant improvements in speech sound production, and it was concluded that the TP model could be an effective and efficient way of providing speech sound intervention to school-age children. On the other hand, the authors pointed out that their small sample size and non-laboratory environment precluded generalisation and that larger scale research trials need to be conducted.

Fairweather, Lincoln and Ramsden (2016) investigated the efficacy of delivering telepractice therapy programmes using low-bandwidth internet to preschool and primary school children (aged 3–12) attending schools in rural New South Wales, Australia. Their sample encompassed 19 children who received therapy via different video conferencing software (Skype, Facetime, and Adobe Connect). As part of their study, they also investigated parents' views, via semi-structured interviews with five parents, of how feasible and acceptable the programme was. In coupling outcome data with parental views, the results indicated that the TP model was feasible and acceptable.

6.11.3 Telepractice applications in stuttering management

The use of telepractice to provide stuttering therapy for different age groups is not uncommon, particularly in countries with relatively small populations that are spread out over a large physical area (e.g. Australia) (Packman & Meredith, 2011). When reviewing the literature concerning TP applications for people who stutter, it consists of 1) anecdotal reports, 2) studies that utilised low-tech means of communication, particularly the telephone, to deliver stuttering therapy services, and 3) studies that utilised high-tech means of communication, particularly video conferencing studies or stand-alone internet therapy programs. In the following sections, the relevant literature is reviewed.

6.11.3.1 Anecdotal reports

In 2000, Kully produced a case report describing the experience of providing a follow-up service for an adult client using dedicated video conferencing facilities. Prior to commencing the follow-up sessions, the client received intensive residential treatment at the Institute for Stuttering Treatment Research (ISTAR) in Edmonton, Canada. During the follow-up period, the client had regular sessions via telepractice that decreased as time progressed, while reporting his satisfaction with the TP method. This study is regarded as the first to report that video conferencing is adequate for deciding the presence or absence of a stuttering moment, although the report was anecdotal in nature and no data was recorded to support the study (Lowe, O'Brian & Onslow, 2013).

Later, Kully (2002) also examined the use of telepractice at ISTAR; this time, with more cases and a wider age range, Kully (2002) reported that over 80 TP sessions had been administered to clients ranging from 3 to 38 years old, with the main motive for adopting telepractice as a service delivery model being the unequal distribution of specialised services. The majority of case treatments were conducted by combining clinic-based treatment with the TP model. Clients were required to travel to telepractice locations, which were either schools or community centres that were equipped appropriately.

Kully (2002) reported that clients and their families were satisfied with telepractice, as they thought that, when compared to travelling to Edmonton to receive treatment, it saved time and money. Certain challenges were apparent, however, stemming from the fact that children were not physically present in the clinic room. For instance, SLTs found it difficult to present resources such as games or activities; instead, families used materials from their homes during the session. SLTs were also not able to directly demonstrate treatment procedures with the child; consequently, SLTs worked indirectly with parents by providing detailed instructions and feedback, specifically with regard to certain problem-solving skills. Both children and adults adapted quickly to the video conferencing system, and building rapport was not a challenge.

Kully (2002) commented extensively on the audio and video quality of sessions. In cases of poor bandwidth, SLTs conducted sessions using audio signals only, as low bandwidth mainly affects visual quality. In this regard, the telephone served as a last-resort backup in case bandwidth proved insufficient. Another technical challenge was the delay in transmission time, for which the SLTs compensated by pausing before responding; however, there

remained a concern that this lag may affect the appraisal of stuttering moments (e.g. subtle silent prolongation). Kully (2002) concluded by raising issues related to the need for client privacy and confidentiality, together with consent, liability, and licensing.

6.11.3.2 Low-tech telepractice applications in stuttering management

Treatment of preschool children who stutter

Three studies have reported on the telephone delivery of treatment for young children who stutter, the first of which was Harrison, Wilson and Onslow (1999), a case study of a five-year-old boy with a history of stuttering who was situated in the UK and was being treated from Australia using the Lidcombe Program. The way the programme is usually delivered was modified to make it amenable for remote use; it involved telephone consultations, along with video and audiotapes delivered by mail, so as to enable the clinician to monitor and evaluate progress. Instead of weekly contact, consultations occurred at a mean rate of 1 every 11.5 days, with reduced times of a median of 45 minutes instead of an hour. Pre-treatment %SS was from 12.4 to 17.7% syllables stuttered. After 25 phone calls over a period of nine months, the subject reached near zero stuttering, which was maintained for 23 months of post-treatment. Harrison, Wilson and Onslow (1999) argued that the subject's stuttering was effectively controlled for a clinically significant time and that, given the subject's age, this could not be attributed to natural recovery. However, it should be noted that 25 telephone consultations were needed for Stage 1 to be completed, which is higher than the usual number conducted in the original programme. Moreover, this study precludes generalisation, as it utilised a single case study design (Lowe, O'Brian & Onslow, 2013).

In an attempt to gather further evidence on applying the Lidcombe Program from a distance, a Phase I trial evidence study was carried out by Wilson, Onslow and Lincoln (2004) using a series of case studies. As in the previous study, Wilson, Onslow and Lincoln (2004) used the telephone to deliver therapy, i.e. although the authors acknowledged the preference of video conferencing over low-tech technologies, infrastructure must be adequate to carry out therapy with such a method, which was not the case at that time. Hence, video conferencing will not improve access to patients who are isolated from regular access to SLT services. Wilson, Onslow and Lincoln (2004) followed the same modifications as those applied in Harrison, Wilson and Onslow (1999), with the addition of video tapes being provided to the parents for

training along with telephone consultations. As in Harrison, Wilson and Onslow (1999), the number of phone consultations required to complete Phase 1 of the therapy (27) was higher than in a clinical setting. For the five children in the study, with an age range of three years and five months to five years and seven months, stuttering pre-treatment was from 2–20 %SS. Data for four children was also available 12 months post-treatment, which was 0.2, 0.6, 1.3, and 1.4%SS, thus suggesting that the method of treatment was viable. However, due to the study design, the extent to which natural recovery contributed to the results is difficult to assess, unlike in randomised early stuttering interventions. It should be noted that the study initially recruited eight families; however, three families dropped out of the study, and this high dropout rate (38%) provides an indication of the acceptability of the TP model when delivering intervention to families residing in rural areas (Lowe, O'Brian & Onslow, 2013).

With regard to a Phase II trial of telepractice delivery of the Lidcombe Program using the telephone, a study was conducted by Lewis et al. (2008), who evaluated the efficacy of TP delivery of the Lidcombe Program using a randomised controlled study design that compared a treatment group with a non-treatment group. As in the previous studies, the programme was delivered using low-tech technology, i.e. the telephone, the justification for which was that more advanced technology was not available at that time. Compared to Wilson, Onslow and Lincoln (2004), the study dropout rate was less (18%). Initially, the study started with nine children in the treatment group and 13 in the control group; however, they ended up with eight in the treatment group and ten in the non-treatment group nine months post-randomisation. The mean %SS was 6.7 for the treatment group and 4.5 for the control group at randomisation. Mean scores were 1.1%SS for the treatment group and 1.9%SS for the control group nine months post-randomisation. As in the previous studies, given that the mean number of telephone consultations ranged from 27–98, this study confirmed that the method of delivery (telephone) required longer than the standard clinical method of the Lidcombe Program. However, a satisfaction survey was distributed to parents, who commented positively on their experience in the study.

As mentioned before, these studies incorporated the use of a low-tech method (i.e. telephone) to deliver treatment. However, this method was believed to be viable at that time given the convenience to SLTs, its cost effectiveness, and the fact that the clients did not have to travel for the sessions. However, using video conferencing to deliver services may make it easier to establish rapport, since the client and the SLT can see each other and the SLT can observe

subtle and covert behaviour that may be not evident when using audio alone (Lowe, O'Brian & Onslow, 2013). Moreover, it facilitates live measurement for observing, coaching, and modelling therapy techniques (Wilson et al., 2004).

Treatment of adults who stutter

O'Brian, Packman and Onslow (2008) conducted a Phase I trial of delivering the Camperdown Program via telephone (for more details about CP therapy, see Section 2.4.3 in Chapter 2). Participants comprised ten adults who stutter, and results at a group level revealed 82%SS reduction immediately after treatment and 72%SS reduction six months post-treatment, with mean telephone contact averaging eight hours (which is less than the original CP). It was concluded that the TP model may be a viable alternative to face-to-face delivery, with relatively less contact with the SLT. In addition, none of the participants dropped out of the programme during the study, although some went through illness, changed jobs, and moved from one city to another. However, it was cautioned that long-term fluency results were only achieved by four out of ten of the participants in this study.

Carey et al. (2010a) conducted a Phase II RCT to determine how efficient this service delivery method was, in which 40 participants were randomised into two groups, the first receiving standard CP in the clinic and the second experimental group receiving CP via telephone. It was found that both clinic and telephone delivery were efficacious. In terms of the time required to deliver therapy, it was observed that participants who received CP via telephone required a mean of two hours less compared to the standard face-to-face group, which was estimated to be a 20% saving per participant. These finding illustrate how cost effective this delivery method can be, as it resulted in less treatment charges for the client and saved clinical hours and resources for the SLT (Carey et al., 2010; Lowe, O'Brian & Onslow, 2013).

6.11.3.3 High-tech telepractice applications in stuttering management

Video conferencing

Treatment of children who stutter

Moving on to more advanced technology in delivering stuttering therapy, a viability study of stuttering treatment on six patients with an age range of 4–19 was conducted by Sicotte et al.

(2003) using dedicated video conferencing facilities. The service connected a Montreal clinic with a remote tertiary centre in Québec, with treatment consisting of 12 weekly treatment sessions initially, and then a schedule of less frequent sessions during the maintenance stage of the treatment process. The study did not provide treatment details. However, though no specific therapeutic method was identified, the authors pointed out that the types of therapy used were evidenced based.

Treatment outcomes were reported based on %SS reduction. In this regard, pre-treatment stuttering levels ranged from 13 to 36%SS, whereas post-treatment stuttering levels reduced to a range of 2–26 %SS and post-follow-up stuttering levels ranged from 4 to 36%SS. All participants were reported to have reduced stuttering levels immediately post-treatment and to some degree at the follow-up assessment. Overall, SLTs were satisfied with the clinical quality of 81% of the sessions. Sessions that were judged as less satisfactory included times when the speech therapist found it difficult to engage with a child moving around or one who was shy. The technical quality was reported to be moderately good, with image reported as being the least successful technical aspect. Participant opinion was obtained using questionnaires, and technical and clinical quality were highly rated by five of the six participants. All participants reported confidence in the quality of treatment. In conclusion, Sicotte et al. (2003) reported that their study supported the use of video conferencing as a viable stand-alone service delivery method for children who stutter. It was concluded that full assessment and treatment of stuttering in children and adolescents can be carried out successfully via telepractice. However, no data regarding assessment was presented.

A study by O'Brian et al. (2014), conducted as a Phase I clinical trial to evaluate the delivery of the Lidcombe Program using video conferencing, specifically via home computers equipped with internet access and webcams, involved three preschool students, who took part alongside their parents. The SLT did not work directly with the children when teaching parents to do verbal contingencies, as in the standard LP therapy practice; instead, the SLT used description and parent roleplay. Results indicated that, from pre-treatment to six months, post-Stage 2 of the Lidcombe Program, Participant 1 reduced stuttering by 81%, Participant 2 by 99%, and Participant 3 by 69%. Parents commented that telepractice reduced travel and cost commitments and caused no interruption to the daily routines of the affected child's siblings during treatment because of travel to and from the clinic. It is worth noting that these children received a stuttering assessment by an SLT via webcam prior to commencing the

study. However, the reliability and validity of delivering assessments via such a method was not evaluated.

A Phase III randomised controlled trial comparing the standard Lidcombe Program treatment and the TP Lidcombe Program treatment was conducted by Bridgman et al. (2016) in order to determine both efficiency and efficacy. Participants were 49 children aged from 3 years old to 5 and 11 months at the start of treatment. The TP group received treatment via Skype. Primary outcomes were the percentage of syllables stuttered at nine months post-randomisation and the number of consultations to complete Stage 1 of the Lidcombe Program. Results revealed insignificant difference between the two groups regarding the %SS post-treatment and parent-reported stuttering severity, as well as insignificant difference in rapport-building between the SLT and parents and the SLT and the children. When the parents of children in the TP group were asked for their feedback regarding the TP model, most parents commented that it was convenient. The authors concluded that conducting the LP via telepractice is viable and efficacious.

Valentine (2014) assessed outcomes in stuttering intervention across three service-delivery models, i.e. the conventional face-to-face model, the TP model, and a combination of the face-to-face and TP model, or a 'hybrid' model, for two 11-year old children who stutter. The aim was to investigate the maintenance of short-term goals via telepractice sessions. Outcome measures were the SSI-IVs administered to each child before and after each intervention period, the %SS calculated from weekly fluency samples obtained in each of the ten-week intervention periods, and the administration of the CAT to assess the children's attitudes toward speaking. In addition, children and their parents completed a questionnaire regarding their experience with TP therapy. Results revealed that both children demonstrated improved fluency measured by weekly %SS, SSI-IV severity ratings showed improvement in one child and remained consistent for the other, and parents and children had positive experiences with intervention via TP; thus, the author concluded that telepractice may be viable for improving and maintaining fluency.

Treatment of adolescents and adults who stutter

Telepractice delivery of the Camperdown Program via webcam has been investigated with adolescents (Carey et al., 2012). Results from a Phase I trial with three participants were positive. As a group, the participants demonstrated an 83% reduction in stuttering frequency

immediately following treatment, a 93% reduction six months post-treatment, and 74% reduction 12 months post-treatment. Further, the participants expressed their preference for telepractice when compared to the clinic-based sessions they had experienced in the past.

A case report by Irani and Gabel (2011) reported that telepractice is effective, especially in the maintenance phase of stuttering intervention. The participant in the study was a 21-year-old male with a long history of stuttering, who was enrolled in an intensive three-week therapy programme at a university clinic. Following the three-week programme, a 12-month maintenance phase was established, which included two weekly sessions for six months and then one weekly session for six additional months. All sessions were conducted using Skype. Outcomes were obtained at the start and end of the intensive programme and during the two six-month interval periods of the maintenance phase. Measures included %SS during various speaking situations, stuttering severity as measured by the SSI-III, and the OASES to measure attitudes and emotions. %SS across various speaking situations, SSI-III, and the OASES decreased dramatically following the intensive programme and continued to decrease for six months post-treatment; however, %SS minimally increased one year post-treatment. It was acknowledged that the study had drawbacks, such as the study design (single subject), in addition to technical breakdowns of the internet, such as delays in or complete loss of audio, occasional freezing of the video stream, and an inability to practice transfer activities via telepractice. However, it was concluded that the TP model was cost effective and efficient for intervention and maintenance.

Stand-alone internet therapy programme application

Treatment of children who stutter

A recent Phase I trial of Internet-LP, with the aim of trialling parent training, was undertaken by Van Eerdenbrugh et al. (2018). The authors reported that the construction of the Internet-LP was carried out keeping in mind that it should resemble the standard LP as much as possible and that parents are the primary users of the program, given that they are required to enter data regularly and should receive individualised feedback from the programme. The authors also mentioned that the Internet-LP was made up of two parts (unlike the standard LP), whereby Part 1 comprised parent training, which was the focus of this paper, and Part 2 guided the parents through the rest of the treatment process after being fully trained. Eight parents were recruited for the study and received training; after completion, the parents were

assessed via the completion of three tasks: the first comprising questions designed to assess their knowledge of stuttering and the LP; the second asking parents to identify their children's stuttering in real-time; and the third comprising 15 open-ended questions about parents' experiences of receiving Internet-LP training. The results revealed that the parents scored well for identifying and measuring stuttering and for knowledge about conducting practice sessions, including how to present verbal contingencies during practice sessions, and the authors pointed out that this trial led to adjustments and that Part 2 would be carried out. The authors also mentioned that the main reason to develop this model was to increase access to treatment for families who are unable to receive it or have limited access to face-to-face contact with an SLT. However, they also noted that not all parents and children suit this service delivery model and that some may require SLT input.

Treatment of adults who stutter

A pilot study (Erickson et al., 2012) investigated the first stand-alone (clinician-free) internet speech restructuring programme for adults who stutter. In particular, two participants completed a computerised adaptation of the Camperdown Program delivered through the internet. Participants reduced their stuttering frequency by 59% and 61% after using the programme over a period of four weeks, in addition to a reduction in self-reported SR and situation avoidance. Results suggested the potential to reduce stuttering without any clinician participation. Further, participants had the convenience and flexibility of accessing treatment at any time and without visiting a clinic, thus reducing the costs associated with clinic fees, travel, and time away from work. The programme also allowed participants to complete the activities at their own pace. Later, a Phase I trial was reported, with results indicating that the study provided further support for the potential of stand-alone internet programmes for adults who stutter (Erickson et al., 2016).

A recent review by Lowe, O'Brian and Onslow (2013) provided a detailed overview of the literature available for the application of telepractice to stuttering management. The studies consisted of delivering stuttering treatment for young children, adolescents, and adults who stutter, most of which have been covered in the previous sections (in addition to studies published after the review). Strikingly, it was pointed out that, when evaluating the body of TP stuttering management research, it entirely lacks, even with basic research methods, any evaluation of the application of assessment methods via telepractice. As not a single study has investigated the viability of TP stuttering assessment with any age group, particularly with

regard to the ability of an SLT to use video conferencing to consistently and accurately identify stuttering behaviours of both an overt and covert nature.

The authors also pointed out that there is a lack of evidence to support incorporating TP delivery models to the stuttering management of school-age children and adolescents who stutter, which may help in increasing access to specialist SLT services for these age groups, given the challenges in providing services for this age group, i.e. they have limited options for treatment due to a lack of SLT specialists with the required experience to deal with such age groups, coupled with increased demands from school, family, and lifestyle commitments, thus making accessing already limited services more difficult. In this regard, the authors encouraged the use of the TP service delivery model to adolescents, as this age group has had a high level of exposure to technology from an early age, in addition to it also being part of their daily lives; moreover, telepractice may increase compliance and motivation to receive therapy as it provides school-age children and adolescents who stutter with a sense of independence, with no need for their parents to accompany them to therapy sessions (Coleman, 1999; Lowe, O'Brian & Onslow, 2013).

The authors also highlighted the need to explore the clinical outcomes of telepractice management in stuttering research, which is achieved by exploring client satisfaction with the service delivery model (Hill & Theodores, 2002; Lowe, O'Brian & Onslow, 2013), in addition to research that aims to identify factors that predict responsiveness to TP applications of stuttering interventions, e.g. stuttering severity and stimulability (Lowe, O'Brian & Onslow, 2013).

In a similar vein, studies have examined alternative methods of gathering speech samples, e.g. investigating the use of telephone calls as a suitable method for data collection (O'Brian et al., 2010). Vogel et al. (2015) investigated the feasibility of using interactive voice response (IVR) technology to collect automated speech samples from ten six-year-old children who stutter. IVR systems are computer-automated telephone technologies commonly used by corporations to interact with their customers. In the study, speech samples were gathered simultaneously by the IVR system and a digital video recorder during three tasks (conversation, picture description, and playing a game). Ratings were compared through the quantification of relative reliability using the ICC, number of syllables spoken, number of syllables stuttered, %SS, and SR. Results revealed that the relative reliability for the video- and telephone-acquired samples on all outcome measures during the conversation task was

high. However, findings were less consistent for speech samples during picture descriptions and games. It should be noted that no measures of absolute reliability were incorporated, and thus agreement between the two methods was not established. In addition, a drawback of the IVR is that it only provides audio information about children's stuttering (Vogel et al., 2015).

As mentioned earlier, Lowe, O'Brian and Onslow's (2013) review indicated that no studies had investigated the viability of conducing a stuttering assessment; however, what is more shocking is that, to our knowledge, since the review was published, there are still no studies, for any age group, regarding the assessment of stuttering via telepractice. Therefore, research is needed to document and establish the clinical value of TP assessment of stuttering. In this context, the main objective of this thesis is to investigate the feasibility of the assessment of school-age children who stutter via telepractice. The main research problem, aims, questions, and hypotheses are presented next.

6.12 Research Problem and Aims

In Saudi Arabia, regular healthcare provisions, including speech and language therapy services, are not available to many of the country's children who live in either major cities or rural and remote areas (AlKabba et al., 2012). Within the current context of accelerated technological advances and the possibility of providing electronic devices that can offer interactive, real-time audio and video communication, one way to improve access for those who may otherwise be deprived of speech and language therapy assessment or treatment is to incorporate telepractice, which refers to speech therapy delivered from a distance via telecommunication technologies. However, within the field of stuttering, very little is known about the reliability and validity of assessing CWS through this modality. Further, the acceptability of this service model among Saudis is unknown. Before it can be implemented, it is of great importance to ensure that assessment results carried out using the TP model are comparable to those performed in a face-to-face environment. Therefore, the researcher intends to conduct a study that will address the issues related to telepractice assessment for school-age children (6-16 years) who stutter. In this regard, such an investigation aims to:

1- Determine the feasibility of the chosen telepractice system for the assessment of stuttering via the internet.

2- Investigate the validity and reliability of telepractice stuttering assessments for children in a clinical setting.

3- Explore parents' and children's views before and after the assessment and determine child and caregiver satisfaction levels with the telepractice assessment of stuttering.

6.13 Research Questions

1- Are telepractice stuttering assessments of children using the selected telepractice system feasible and comparable to face-to-face assessment?

2- Will the telepractice stuttering assessment of children yield results that are comparable to a face-to-face assessment?

3- How satisfied will parents and children be with the telepractice service provided? Are there any changes in the children's and their caregivers' views before and after conducting the assessment?

6.14 Research Hypotheses

H1: Using the selected telepractice system, the telepractice stuttering assessment will be feasible and comparable to the face-to-face assessment.

H2: Results from the assessment will present a high degree of clinically acceptable levels of agreement and reliability between the two environments, validating the telepractice modality for stuttering assessment.

H3: Children and their parents will feel positive about the potential of using telepractice for stuttering assessment, satisfied with the telepractice assessment of their stuttering, and willing to use the telepractice system for stuttering therapy in the future.

7 Methodology

7.1 Introduction

The goal of this chapter is to provide a description of the research design adopted and methodology used to tackle the research questions. In this regard, the following sections are dedicated to explaining the study design in the method comparison study. This chapter also includes inclusion and exclusion criteria, materials and equipment used, ethical considerations, process of recruitment, and allocation of testing environments, testing, and scoring. Finally, the chapter provides an explanation of the data analysis strategy.

7.2 Methods

7.2.1 Research design

The present research study implemented a method comparison design (Bland & Altman, 2003). This design was chosen as it has the capacity to assess the variation between two methods for the same subject (Sutherland et al., 2017); more specifically a non-inferiority methodology was utilised, as the question being posed is whether a TP stuttering assessment can provide similar outcomes to a face-to-face stuttering assessment. The aim is not to replace the conventional face-to-face SLT assessment service with the telepractice service but rather to offer a TP service option for people who cannot access services (Russell et al., 2017). To ensure the reliability of findings, agreement between two SLT scorings made in the two environments (face-to-face and at a distance) was measured simultaneously. Such simultaneous scoring offers the opportunity to compare ratings obtained in both assessment environments while lessening the test-retest or learning effects, which may occur if a child was assessed online and face-to-face in a serial order on two occasions (Waite et al., 2010). Also, if participants were assessed serially, this could create data error due to participants' variability between assessments (Hill et al., 2006). In our study, the target population is CWS, and it is well documented that stuttering behaviours can vary significantly across situations, time, and listeners (Alameer, Meteyard & Ward., 2017). Therefore, the serial assessment method may be cumbersome as participants could perform differently in both situations, resulting in two different ratings across environments, i.e. related to the variation in the participant's state rather than the method of assessment (Russell et al., 2017). The simultaneous scoring procedure has been incorporated in several TP assessment studies in the field of SLT (e.g. Hill et al., 2009; Palsbo, 2007; Theodoros et al., 2008; Waite et al., 2006;

Waite et al., 2010). An important aspect that should be investigated when conducting a method comparison trial in the area of telepractice is the reliability of the proposed new TP assessment (Russell et al., 2017). Hence, the reliability of the new stuttering TP assessment needed to be established. Therefore, 30% (n=10) of the telepractice assessments were re-rated, and intra- and inter-rater reliability were determined. In addition, the study also established the reliability of face-to-face assessments.

7.2.2 Participants

The study used a convenience sample. In this regard, participants attending the Speech and Language Therapy Clinic at King Abdulaziz University Hospital (KAUH) in Saudi Arabia, particularly in Riyadh, who met the inclusion and exclusion criteria were recruited. In total, 30 CWS were recruited. Two assessors participated in this study, the main researcher and another SLT. The principal researcher and another SLT were qualified and had experience of one year (principal investigator) and ten years (second assessor) in the assessment and therapy of different domains of paediatric speech and language disorders, including stuttering.

7.2.2.1 Inclusion criteria

- Participants either with identified stuttering or newly referred with suspected stuttering, aged between 6 and 16.
- Arabic as their first language, as language differences may serve as a confounding variable in the study.

7.2.2.2 Exclusion criteria

- Participants with a significant hearing, neurological, or cognitive impairment or uncorrected visual impairment.
- Other primary diagnoses (e.g. Down syndrome, ADHD, SLI, and dyspraxia), based on parental report.
- First language is not Arabic

7.2.2.3 Recruitment

The main researcher approached potential participants following the criteria mentioned above. The researcher explained the study to both parents and the child and handed them an information sheet (see Appendix 4). If the parents and child agreed to participate, the child was scheduled for assessment.

7.2.3 Ethical considerations

7.2.3.1 Ethics

Prior to commencement of the study, approval from the University of Reading Ethics Committee was sought. After receiving approval, the next step was to apply to the Institutional Review Board (IRB) of the School of Medicine at King Saud University (see Appendix 5). After gaining approval, approval from KAUH was obtained, following Saudi guidelines by contacting the head of the SLT department. The letter sent seeking permission to conduct the study at the out-patient clinic detailed the allocation of the patients' process after obtaining their informed consent. The final step was engaging SLTs to recruit participants and, for the researcher, another SLT to carry out the assessments and fill out the questionnaires.

7.2.3.2 Anonymity, confidentiality, and data protection

According to Data Protection Act guidelines (Sauerwein & Linnemann, 2001), participants were coded using numbers to anonymise their identities in all data collected during the TP assessment or face-to-face sessions. Video recordings were also coded with this research code. Similarly, hard copies of the assessment forms used by both SLTs involved in the assessment were stored in a locked filing cabinet in a locked office. All electronic data and analyses were stored in a password-protected computer, which was only accessible to the principal researcher and the supervisors of this research.

7.2.3.3 Consent

Consent was obtained in writing from both parents or guardians and from the child if possible. If it was not possible to obtain written consent from the child, it was obtained from the guardian and verbal consent was obtained from the child (see Appendices 6 and 7). By signing the consent form, the guardians also consented to using audio-visual and photographic material for later viewing by the lead researcher and SLTs involved. Participants were informed that they could withdraw from the research at any time without consequence or compromise to their stuttering management at the hospital.

7.2.4 Materials

The materials and instruments for the current study included (a) materials and tools used for the stuttering assessment (see Section 7.2.4.1) and (b) materials needed for video conferencing (see Section 7.2.4.2).

7.2.4.1 Assessment tools

SSI-IV

The Stuttering Severity Instrument (SSI-IV) was chosen, as it is the most widely used syllable-based procedure for assessing behavioural symptoms of stuttering (Riley, 1994; Riley, 2009). The total score is calculated based on the sum of frequency, duration, and physical concomitant behaviours. For more details on applying and scoring the SSI-IV, see Appendix 2.

Test procedure

The test, including all its elements, was administered in accordance with the testing procedure set out in the examiner's manual highlighted in Appendix 2. In both administration conditions, the participants' responses were recorded via video by the face-to-face SLT and the remote SLT. Both SLTs scored responses from the recorded material after the session.

Each participant was assessed using the Stuttering Severity Instrument SSI-IV. Pictures and reading material were digitised for this study. In the TP format, picture and reading plates were presented through the video conferencing system, while, in the face-to-face format, materials were presented in printed form.

Percentage of stuttered syllables (%SS)
that
%SS is a measure of the proportion of syllables in a speech sample that contains unambiguous stuttering, and it is known as the gold standard measure for determining stuttering severity (Jones et al., 2005) (for more details regarding this measure, see Chapter 2). %SS is also part of the SSI-IV, which represents the raw score of %SS in the test, calculated using the following formula:

total number of stuttered syllables ÷ total number of syllables spoken x 100

To score the %SS, the same procedures highlighted in the %SS section in the SSI-IV were followed.

Severity Rating scale (SR)

The Severity Rating scale is a perceptual measure that is simple to use and requires no equipment (Onslow, 2017) (for more details regarding this measure, see Chapter 2). It is, by concept, different from %SS, where the rater is not required to count stuttering moments but assigns a numerical value that represents perceived overall stuttering severity (O'Brian et al., 2004). Perception of severity is recorded on a scale from 1–9, where 1=no stuttering, 2=extremely mild stuttering, and 9=extremely severe stuttering. When assigning a number to the sample heard, the rater takes into consideration stuttering rate, stuttering frequency, and the severity of individual moments of stuttering (O'Brian et al., 2004).

7.2.4.2 Equipment for video conferencing

Video conferencing was conducted using Cisco WebEx (http://www.**webex**.com). During the search and selection period for an appropriate video conferencing system, the researcher kept in mind the criteria for selecting video conferencing software that have been reported in telepractice research (Matthews, 2014):

- To select a system that is simple to use
- Can be used in the location of the child
- Does not need on-hand engineer support
- Requires little or no additional equipment
- Low cost or free for the client.

Initially, three video conferencing software that met such criteria were considered: Skype, GoToMeeting, and WebEx. However, Skype and GoToMeeting were disregarded, as Skype was not permitted for use in Saudi Arabia at that time and, after trialling GoToMeeting, the audio and image quality were not satisfactory; in this regard, WebEx was superior and was therefore chosen for this study.

The WebEx system can be downloaded on any computer, laptop, or mobile phone. It uses the built-in webcams and microphones found in computers and is free to download. It is also a HIPPA compliant video conferencing system. Moreover, additional features are as follows (Walker, 2015):

- ***File, application, and desktop sharing:*** Users can easily collaborate during sessions while sharing audio and video.

- ***Comprehensive multimedia experience:*** Users can share video files in real time and incorporate multimedia in sessions.
- ***Real-time video for several participants:*** WebEx provides multiple, simultaneous webcam video feeds with voice-activated switching. Up to 25 people can view and/or participate in the session in real-time.
- ***High-definition video:*** The HD video, with up to 720p screen resolution, provides exceptionally clear images of the people and materials.

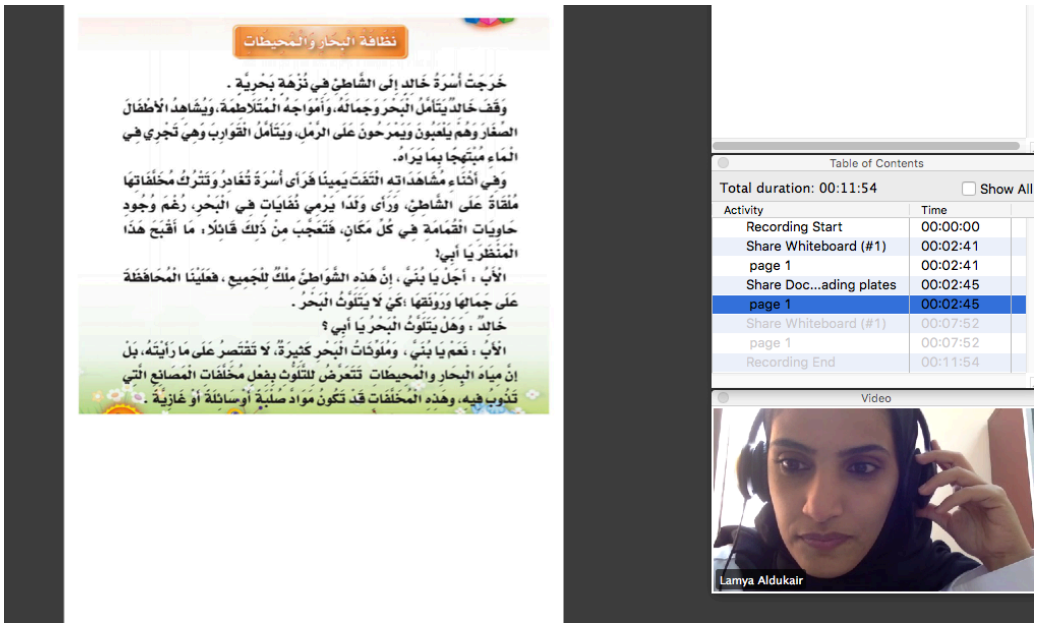


Figure 2 Screenshot of Reading Passage Presented and Shared Online via WebEx Whiteboard from the Participant's End

Equipment consisted of two laptops and their built-in cameras and microphones (one 15-inch MacBook Pro (at the child's end) and a 13-inch MacBook Air (at the distant therapist's end). Both laptops had a built-in HD webcam, microphone, and speakers. The 15-inch MacBook Pro, which was at the child's end, had a 2.2 GHz processor with 16 GB memory. During TP sessions, the 15-inch screen was large enough to clearly present assessment material and view the clinician during video conferencing. The 13-inch MacBook Air, at the distant clinician's end, had a 1.4 GHz processor with 4 GB memory. During TP sessions, the screen was large enough to observe and interact with the child when the TP SLT was leading, and, during face-to-face sessions, the screen was large enough for the TP SLT to observe the child silently when the face-to-face SLT was leading. For all conditions, the built-in laptop webcam was

used to view the sessions, all of which were recorded using the video recording feature of WebEx web-based software for the telepractice condition, in addition to Many Cams (a free desktop capture software) for face-to-face scoring. For the TP sessions, all materials used for the assessment were electronic and were presented using the whiteboard feature in WebEx (see Figure 2 and Table 10). Internet connectivity was established using two dongles, one in each site, to connect to the Wi-Fi.

Table 10 Equipment Description and Use

Equipment	*Description*	*Location, Condition, and Use*
 	Laptop 1: -13-inch MacBook Air - Built-in HD webcam, microphone, and speakers. -1.4 GHz processor - 4 GB memory Laptop 2: -15-inch MacBook Pro -Built-in HD webcam, microphone, and speakers. - 2.2 GHz processor -16 GB memory	Laptop 1: Was used by the distant SLT for video conferencing to connect with the participant (SLT to participant). Laptop 2: Was used at the child's end for video conferencing to connect with the SLT (participant to SLT).
	- V7 deluxe stereo headset, black - Wired: 3.5 mm plug	Was used by the distant clinician for video conferencing.
Cisco webex	Video conferencing software	Was used in both sites. However, the distant clinician was controlling it. The software was used for video conferencing in both sites. In addition, it was used to record the session by the distant clinician.

	Huawei E5573s mini Wi-Fi dongle	Two dongles were used, one for each site (TP and face-to-face) to connect to the internet.
Manycam	Free webcam software	This application was used to record sessions by the face-to-face SLT

7.3 Procedure

7.3.1 Pre-study clinician training

In order to ensure consistency in stuttering identification and data recording, both the main researcher and the second SLT were trained on scoring three samples prior to the start of the study. Reliability was not formally established; however, the clinicians sat in a room together, listened to the samples, and then rated the samples and provided scores of %SS, SR, and SSI-IV independently. After that, scores of each measure were compared, in case the scores of any of the three measures differed. Samples were then replayed and any discrepancies discussed until consensus on scores was reached. This step was taken to make sure that both clinicians agreed on the same criteria of what did and did not constitute a stuttering moment, as this could have been a problem as disagreement or different perspectives about what constitutes a stuttering moment may serve as a confounding variable.

For the TP system, prior to the commencement of the study, the main researcher trained the second SLT on the use of the video conferencing system and all the roles that the second SLT might take during the study (i.e. leading the distant SLT, observing the distant SLT, leading

the face-to-face SLT, and observing the face-to-face SLT). The SLT was also trained on how to login to the video conferencing software, start a video conference call, position the camera, set up the room and lighting, and troubleshoot. The training took two sessions, lasting for an hour each.

7.3.2 Blinding

In an ideal situation, it is recommended that participants, SLTs, and whoever is involved in the data collection process are blinded to the randomised condition in order to avoid bias in performance during the session or scoring. However, this was an interactive telepractice study, and it was difficult to blind SLTs and participants to the assessment conditions (Nelson & Palsbo, 2006). Moreover, although samples would be scored later, raters of the sample were the SLTs who carried out the assessment, and therefore blinding of scoring was not possible. However, the third SLT who did not take part in this study and was responsible for rating part of the samples for reliability was blinded to the purpose of the study.

7.3.3 Assessment procedure

The assessment took place in two rooms within King Abdulaziz University Hospital, Riyadh, with the distant SLT in one room and the face-to-face SLT with the child in the other. For each assessment, two SLTs were randomly assigned their role (the main researcher and another SLT), and the participants were randomly assigned to an administration type, either TP-led or face-to-face led assessment. In total, 15 online-led and 15 face-to-face-led assessments were conducted. For each session, either the online SLT or the face-to-face SLT administered the assessment while the other SLT acted as an observer. At the beginning of each assessment, the leading SLT interviewed the child, and the parent accompanied the child and filled out a case history form. Following this, the rest of the assessment tasks were carried out.

7.3.3.1 Case history and interviews

The case histories used for the child interview questions were those from the Stuttering Clinic at King Abdulaziz University Hospital. The leading SLT and parent engaged in a discussion regarding the child's general case history. The parents were asked about the general development of the child, onset of the stutter, the child's environment, including school and family dynamics, and the presence of any parental concern. Moreover, questions also targeted the child's self-perception and awareness and the impact of stuttering on daily life.

After that, the SLT engaged in a discussion with the child, who was asked general questions about his/her stuttering. Specifically, the clinician asked questions about the child's feelings and attitudes toward his/her stutter in different situations, e.g. with strangers, at school, and at home. In addition, questions addressing avoidance or escape behaviours were included. If the SLT felt that discussion of the child's case should be done with the parents alone and without the child present, the second SLT accompanied the child to a nearby room until the leading SLT was finished.

It should be noted that the information gathered from the interviews and case history was not included in the analysis. However, the goal of carrying out interviews with children and their parents was to mirror a typical stuttering assessment, therefore providing them with the full experience of receiving a stuttering assessment in order to be able to provide their views and evaluate their satisfaction.

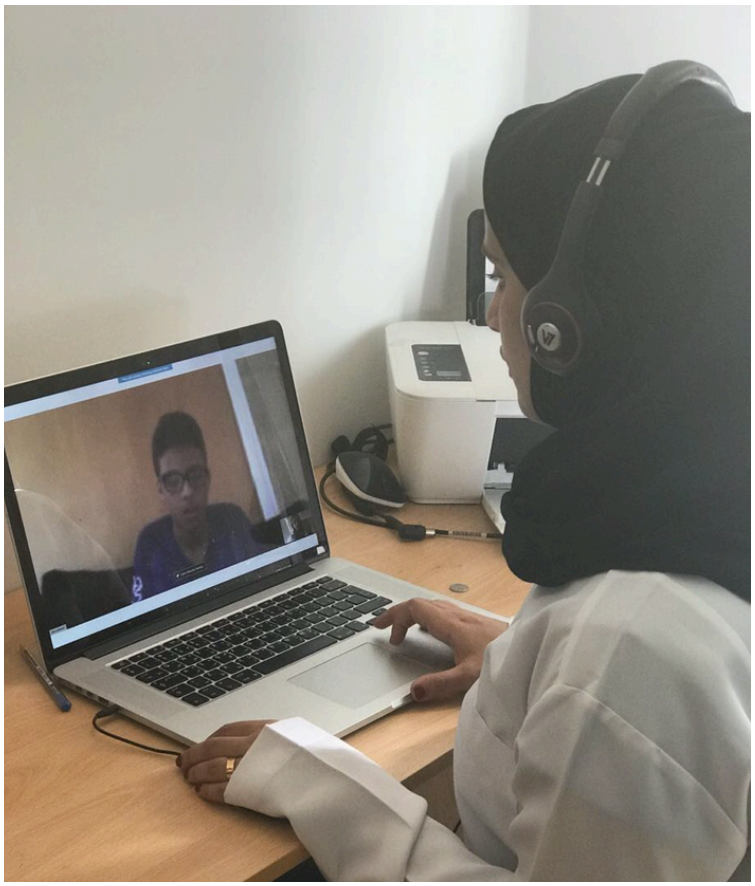


Figure 3 Picture of a TP-Led Assessment from the Leading SLT Site

7.3.3.2 Telepractice-led assessment procedure

The TP SLT led each session and assessed children from a distance via the telepractice system. The child was seated and positioned in front of the laptop camera so as to offer a clear view for the TP SLT, who established the video conference via the system by sending a WebEx link invitation by email. The face-to-face SLT acted as an observer and a facilitator of the assessment by, for example, helping to establish the video conference call via WebEx and repositioning the child if he/she moved from the view of the camera. Sometimes, if needed, the face-to-face clinician interacted with the child upon the distant SLT's request. The distant clinician recorded the session via a telepractice WebEx recorder, while the face-to-face clinician recorded the session via a laptop recorder, which is similar to the typical digital recorder used in face-to-face sessions (see Figure 3).

7.3.3.3 Face-to-face-led assessment procedure

This assessment was carried out following a typical clinical practice of stuttering assessment. The child and face-to-face SLT were positioned at an angle that offered the distant clinician (who was an observer in this testing mode) a clear view of the child. The face-to-face SLT used printed material for assessment that is used in standard practice, in addition to leading the session and assessing the child. As for the online-led assessment, the face-to-face SLT recorded the entire assessment on the laptop to score later, and, while the assessment was being carried out, the online SLT recorded the session via WebEx recorder for the same purpose.

7.3.4 Speech samples

Recordings of the 30 sessions were made, with each participant having two speech samples. Where the child was a reader, one spontaneous sample and one reading sample was elicited, while, for non-readers, two spontaneous samples in two situations were elicited (one sample was a conversation with the clinician and the other was a conversation with the parent with the leading SLT observing). Topics of these spontaneous samples were those familiar to children, e.g. hobbies, school, favourite TV shows, and holidays. As per SSI-IV procedure, picture plates were presented to children to help elicit conversation. In total, 56 samples were elicited, where four non-reader participants recorded one sample only. The length of recordings varied depending on the stimulability of the child. The number of total syllables

was more than 200, thus following the recommended guidelines (Riley, 2009). In this regard, samples ranged from 245 to 1,113 syllables, with a mean of 495.

7.3.5 Scoring of speech samples

The two SLTs who took part in the assessments rated the samples, measuring SSI-IV, %SS, and SR for all 56 speech samples. SSI-IV and %SS were scored together, whereas the SR was scored three months later. Two months after their last measurement task, the main investigator and a third SLT who did not take part in the assessments, who was a qualified SLT with one year experience of different domains of paediatric SLT disorders including stuttering, assessed 30% of samples in both environments (ten samples of the face-to-face-led assessments to measure %SS and SR, as face-to-face SSI-IV reliability data was measured in a separate investigation (see Chapter 5), and ten samples of the telepractice-led assessments of SSI-IV, %SS, and SR for reliability testing). This period of two months was chosen to lessen learning effects and familiarity (Waite et al., 2010). Similarly, %SS and SSI-IV were scored together, whereas SR was scored three months afterwards. The third SLT was blinded to the purpose of the study. SLTs were required to count stuttering moments (as per SSI-IV criteria, see Appendix 2) using dots and slashes, in addition to measuring SR using a 9-point scale (see Section 7.2.4.1.3). To maintain consistency among scoring procedures, raters were only allowed to view recordings three times. In total, there were four scoring conditions: online scoring in the online-led assessment, online scoring in the face-to-face-led assessment, face-to-face scoring in the online-led assessment, and face-to-face scoring in the face-to-face-led assessment.

7.3.6 Questionnaire: Perceptions and satisfaction

To analyse satisfaction regarding the telepractice assessment, questionnaires exploring children and parents'/caregivers' perceptions were given out immediately before and after the TP-led assessment. Pre- and post-assessment questionnaires consisted of the same items; however, the pre-assessment questionnaire was in the future tense, whereas the questionnaire after the session was in the past tense. Both questionnaires explored perceptions regarding how comfortable the child and the caregiver were with the TP audio and video clarity, in addition to general questions about the TP session. Responses to questions in both questionnaires were scored on a 5-point Likert scale (1=strongly disagree, 2=disagree, 3=I don't know/neutral, 4=agree, and 5=strongly agree), in addition to a short answer question. The questionnaire was designed for this study, the methodology was adapted from Sharma et

al. (2013), and questions were taken from a range of sources (Carey et al., 2010; Grogan-Johnson; 2012; Lewis et al., 2008; LoPresti, Jinks & Simpson, 2015; Sharma et al., 2013; see Appendix 8). The questionnaires were completed by the main researcher transcribing the parents' and children's responses to the 5-point Likert scale questions and the short answer question.

7.4 Data Analysis

7.4.1 Research variables

For this investigation, there was one independent variable (IV), which was the method of service delivery (telepractice vs. face-to-face).

The dependent variables (DV) were as follows:

1. The number of sessions completed with success.

2. Telepractice and face-to-face ratings of visual and audio quality.

3. Stuttering assessment scores using different measures (SSI-IV, SR, and %SS) rated in the telepractice and face-to-face environments.

4. Satisfaction ratings of the children and their parents who underwent a telepractice assessment before and after to provide a sense of the acceptability of this model.

7.4.2 Measures

In summary, measures sought in this study were as follows:

7.4.2.1 Measures of feasibility

The TP clinician reported the number of sessions that were completed with success and recorded any issue that precluded the completion of a session and what action was taken to overcome any difficulty encountered. Difficulties included any factor related to the child assessed, the clinician, or the technology. Moreover, the TP clinician separately rated the audio and video quality on a 5-point scale (poor=1, fair=2, good=3, very good=4, and excellent=5).

7.4.2.2 Measures of validity and reliability

Measures of validity and reliability encompassed both the TP and face-to-face SLTs, simultaneously scoring the %SS, SSI-IV and its subcomponents (%SS, duration, physical

concomitants, total score, and final severity rating) and SR. A method comparison analysis (see Section 7.2.1) was used to determine the level of agreement and reliability, which asked whether the two environments (face-to-face vs. telepractice) could be used interchangeably without affecting the participants' results. In other words, this method compared the two environments to detect bias. The statistical analysis chosen permitted a comparison of scores between the two assessment environments on every scale chosen and provided an indication of whether the telepractice system chosen influenced the accuracy of the scoring of the online assessments.

Before explaining the statistical analysis employed to answer our research questions, it is important to clarify the difference between agreement and reliability and different statistical analysis tied to these concepts, given the nature of this study and because of the significant confusion and misunderstanding that currently exist relating to the analysis of agreement and reliability of data generally (Hallgren, 2012) and for stuttering measurements specifically (Cordes, 1994).

Agreement and reliability are often used interchangeably in speech and language therapy research, particularly stuttering research, although some studies highlight their differences (Cordes, 1994; Davidow & Scott, 2017; Karimi et al., 2014a). However, agreement and reliability are two different concepts, and a clear distinction between the two terms is important in order to enable us to select appropriate statistical approaches and provide adequate interpretations (Bartlett & Frost, 2008). Reliability refers to the agreement of two measurements in settings where neither one is assumed 'correct', often referred to as consistency, reproducibility, and repeatability (Müller & Büttner, 1994). In this case, the measurement error is related to the variability between study objects and not to the measurement itself. Therefore, reliability statistics offer a description of a group of participants being studied and do not pertain to individual participants (Karimi et al., 2014a). Two forms of reliability exist: (a) inter-rater reliability, which refers to a situation whereby two or more raters apply the same measure on a sample on one occasion and (b) intra-rater reliability, whereby the same rater applies the same measure on the same sample on separate occasions (Karimi et al., 2014a). Agreement can be defined as the closeness of measurements to a standard reference, i.e. a hypothetical true measurement (Jones, Dobson & O'Brian, 2011; Müller & Büttner, 1994). Unlike reliability, agreement is a property of the measurement methods under investigation, which does not depend on the variability between

study objects (de Vet et al., 2006). In the literature, agreement may be referred to as conformity, correctness, accuracy, or validity (Müller & Büttner, 1994). When carrying out clinical research, agreement analysis is helpful for determining temporal changes within an individual (Karimi et al., 2014a), whereas reliability indices are helpful for comparing between groups or subjects (Atkinson & Nevill, 1998).

To answer the question of whether two datasets agree, a common statistical procedure is to calculate a correlation coefficient, namely Pearson's correlation coefficient, which is commonly used in both telepractice and stuttering research. Pearson's correlation ranges from +1 to −1. A correlation of +1 means that there is a perfectly positive linear relationship between variables, a correlation of −1 means that there is a perfectly negative linear relationship between variables, and a correlation of 0 means that there is no linear relationship between the two variables (Giavarina, 2015). The correlation coefficient may thus appear to be a helpful index. However, its use is strongly discouraged when assessing the notion of agreement (Bland & Altman, 2003) because Pearson's correlation coefficient measures linear association and is therefore unable to detect the presence of systematic bias (Bland & Altman, 2003). This implies that two compared datasets may have a high correlation but systematic bias causing one set of scores to be consistently higher or lower than the other (Karimi et al., 2014a). This can be misleading, as there may be a strong correlation but poor agreement between two sets of data (Bland & Altman, 1986). In addition, this index is highly influenced by the variability of the sample—the larger the range of values, the higher the correlation coefficient (Müller & Büttner, 1994). This makes comparisons beyond the sample examined misleading. Any results for the index can only be generalised to samples with similar variations (de Vet et al., 2006). Due to the weaknesses highlighted above, and because a correlation coefficient measures association rather than agreement, its use was disregarded in this study.

Another common way to assess agreement is to conduct a test of difference, particularly a paired t-test or, in the case of non-parametric data, a Wilcoxon signed rank test (Watson & Petrie, 2010). This test detects the presence of systematic difference between two sets of compared data, and one may therefore argue that it is a better option than conducting a correlation coefficient (Watson & Petrie, 2010). However, this measure can only detect systematic bias and not agreement; in this sense, a significant result suggests that there is a systematic difference, but a non-significant result only indicates that there is no evidence of a

systematic effect and does not indicate whether they agree (Zaki et al., 2012). The reason for its failure to assess agreement is that the mean value may be affected by each data point, i.e. the mean could be over- or underestimated when extremely large or small values are present (Zaki et al., 2012). Moreover, it is possible that poor agreement between the two datasets can be hidden in the distribution of differences, and thus the two datasets can appear to agree (Zaki et al., 2012). When we assess agreement, our interest is not in the mean of scores by each method but in each individual score (Zaki et al., 2012). In summary, this method can detect systematic error, but it is not sufficient on its own to determine agreement. In our study, we ran a Wilcoxon signed rank test on all measures intended to be compared; however, we took this as an initial step to determine whether a systematic difference existed between ratings in the face-to-face and TP environments and carried out more statistical analysis to determine agreement and reliability.

Another common method to assess agreement between two datasets is the intraclass correlation coefficient (ICC). The advantage of this test over Pearson's correlation coefficient is that it accounts for the systematic difference between two datasets (Bland & Altman, 1990). It is calculated as the ratio of the inter-subject component of variance to the total variance (inter-subject variance + within-subject variance) (Lee et al., 2011). However, as for Pearson's correlation coefficient, the ICC is affected by the range of values across a population (Van Stralen et al., 2008), due to which ICC is interpreted as a measure that assesses the degree to which subjects differ from, rather than agree with, the population (Barnhart et al., 2007; Vangeneugden et al., 2005). In other words, it provides no information on the extent to which raters agree on closeness to a hypothetical true score for an individual; rather, it provides information on the degree to which they vary between two judges (Karimi et al., 2014a). Therefore, it is the preferred method for assessing reliability as opposed to agreement (de Vet et al., 2006). In our study, we used the ICC as a reliability index and tested the inter-rater and intra-rater reliability of our data in both environments (face-to-face and telepractice) using these statistics. Different ICC formulas can be applied (see Appendix 9) that are dependent on the following: whether each subject is measured by the same rater or different raters, whether raters are selected randomly or if they are the only possible raters from the rater population, and whether the scores obtained are from a group of judges or a single judge (Karimi et al., 2014a). In our study, we used a two-way, random-effect intraclass correlation ICC (2,1) as our aim is to generalise findings to a general population of raters (Koo and Li, 2016). As for the interpretation of the ICC values, we used Munro's taxonomy

(Domholdt, 2005), where a value of .50–.69=moderate correlation, .70–.89=high correlation, and 0.90–1.00=very high correlation.

To measure agreement, we used Bland and Altman's (1986) graphical method, where the difference between the scores of two measurements is plotted against the mean for each subject. In this method, the mean difference between two methods, or what is called the bias, and 95% limits of agreement (LOA) are calculated. The LOA can be calculated as:

$$\text{LOA} = \bar{d} \pm 1.96s$$

where $\bar{d}$ is the mean difference between face-to-face and online clinician ratings and s is the standard deviation of the difference (Bland & Altman, 1986).

The graph is plotted on an XY axis, where X represents the difference between the two measurements and Y represents the mean of the two measurements. It is expected that 95% of the data points should lie within ±1.96 standard deviations (SD) of the mean difference, which is the LOA (see Figure 4) (Bland & Altman, 1986).

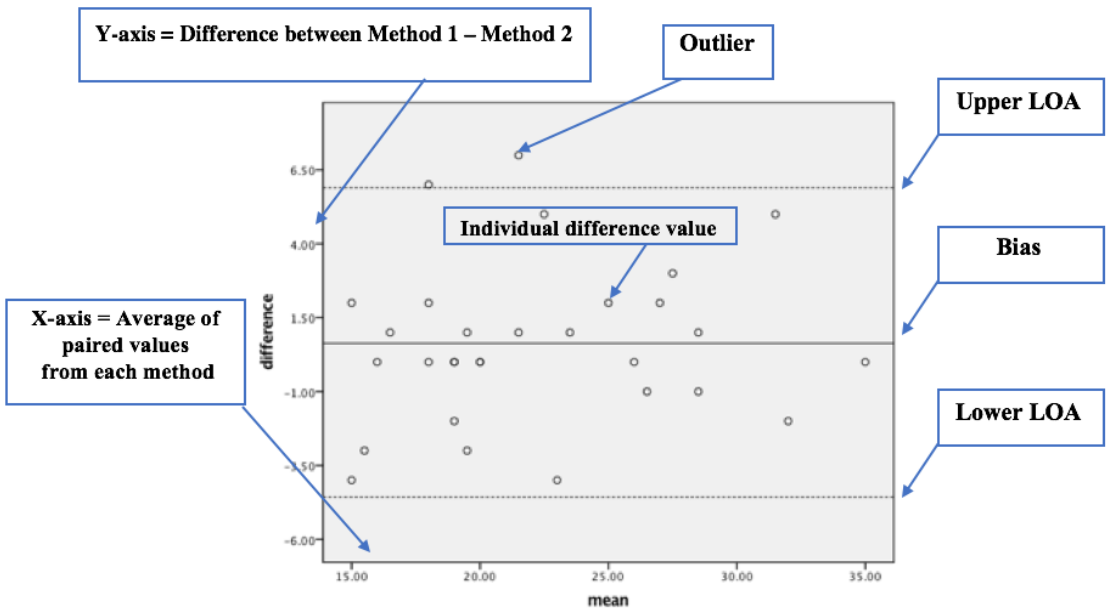


Figure 4 Structure of a Bland-Altman Plot with Explanation of Elements

We applied this method to measure the agreement of all the stuttering measurements chosen (the %SS, SR, and SSI-IV and its subcomponents) between the face-to-face and TP environments. The previous graph shows the difference between face-to-face and online ratings against the mean of the two ratings for each subject (Bland & Altman, 2003). Thus, for the ease of evaluation, it enables visual interpretation of the size and range differences between the two assessment methods (Liaw et al., 2008). However, the acceptable magnitude of the difference is a clinical rather than a statistical decision, i.e. the question of how small the difference is depends on the clinical context: would a difference between online and face-to-face ratings such as that described by the 95% LOA meaningfully affect the interpretation of the results? (Myles & Cui, 2007).

We also assessed agreement using a percentage agreement calculation, which is a simple way to calculate agreement and reliability frequently used in stuttering research (Davidow & Scott, 2017; Karimi et al., 2014a; O'Brian et al., 2004; Teesson et al., 2003). This is derived by calculating the percentage of judge score pairs that differ by zero, 1 point, 2 points, and so on. A limitation of this calculation is that it does not assess judge scores against a hypothetical true measure, in addition to also not accounting for the expected chance agreement between judges (Karimi et al., 2014a).

Despite the given limitation of this procedure, this index was used to calculate agreement of SR and SSI-IV and its subcomponents in order to facilitate comparison of SR and SSI-IV scores with other studies that investigated the reliability of SR and SSI-IV using this analysis.

Finally, all the analyses described above are used to assess the agreement and reliability of quantitative measurements. However, to assess the agreement and reliability of the final severity rating of the SSI-IV, which is a categorical measure, the quadratic weighted Kappa (k_W) statistic was used (Landis & Koch 1977). The k_W is widely used in telepractice studies for ordinal data (Constantinescu et al., 2010; Theodoros, Hill & Russell, 2008; Waite et al., 2010) and provides an indication of agreement between raters (Landis & Koch, 1977). In the present study, the statistics provided a measure of agreement beyond chance between the online and face-to-face measures of stuttering. The k_W assigned weights to the observed and chance agreement and presented levels of agreement, where a k_W of less than 0.20 is interpreted as poor, 0.21–0.40 is fair, 0.41–0.60 is moderate, 0.61–0.80 is good, and 0.81–1.00 is very good (Landis & Koch, 1977).

In summary, the analysis of SR, %SS raw and scaled scores, the three longest stuttering moments scaled scores, physical concomitants, and total scores of the SSI-IV were performed using the Wilcoxon signed rank test and the Bland and Altman LOA method. In this regard, both methods were used together to comprehensively examine the agreement of stuttering assessment in telepractice. This was in addition to the k_W, which was used to measure the agreement and reliability of the final severity rating of the SSI-IV. Due to the problems highlighted in the previous paragraphs, it is essential, in this research context, to have visual representation of the data using the Bland and Altman method, which will allow an examination of differences within subjects (Rankin & Stokes, 1998) and may help us to extract as much information as possible regarding the applicability of the TP model in stuttering assessment.

The intra- and inter-rater reliability of the online and face-to-face assessments were calculated using two-way, random effect intraclass correlation ICC (2,1) for 30% of participants (n=20 participants in total; n=10 participant in each environment) in each environment, except for the overall severity rating, as the k_W was employed. For inter-rater reliability, the third SLP, who did not take part in the main study, rated the videos.

7.4.2.3 Measures of acceptability

The parents and children who participated in the online-led condition completed pre- and post-assessment surveys to measure their acceptance of, and satisfaction with, the model with regard to aspects such as visual and auditory clarity and level of comfort and to gather general information regarding telepractice. Responses were measured on a Likert scale (1=strongly disagree, 3=neutral/I don't know, and 5=strongly agree), in addition to there being open-ended questions. The Wilcoxon test was used to determine if there was a significant difference between pre-and post-assessment views of children and parents.

8 Results

8.1 Introduction

As stated in previous chapters, the purpose of this thesis is to evaluate the application of an internationally recognised stuttering assessment (SSI-IV), along with common stuttering measures (%SS and SR), of school-age CWS via telepractice. In this regard, the thesis' aims are threefold: assessment of the feasibility, validity and reliability, and acceptance of the TP model. Three hypotheses were posed in alignment with these aims, each of which is stated below, along with the results of the data analysis for each question. Before presenting our findings, however, study participants' background characteristics will be delineated.

8.2 Background Characteristics of the Study Participants

Participants in this research comprised 30 school-age CWS, 3 females and 27 males, aged 6 to 15 (M= 9.6; SD= 2.2). The assignment to testing situation (telepractice vs. face-to-face) was randomised, which resulted in 15 children tested in the TP-led situation and 15 children tested in the face-to-face-led situation (see Table 11). Three SLTs were involved in this research. The main investigator and another SLT conducted and rated the assessments of face-to-face and TP sessions, whereas a third SLT was involved in assessing the inter-rater reliability of the TP sessions. In this regard, the main investigator re-rated samples for the intra-rater reliability check of the TP sessions.

Table 11 Background Characteristics of the Study Participants

Face-to-Face-Led			**TP-Led**		
Participant	**Age (years)** **M=8.7** **SD=1.79**	**Gender**	**Participant**	**Age (years)** **M= 10.5** **SD=2.4**	**Gender**
1	10	M	**1**	9	M
2	8	M	**2**	10	M
3	9	M	**3**	9	M
4	7	M	**4**	9	F
5	8	M	**5**	9	M
6	13	M	**6**	12	M
7	8	F	**7**	13	M
8	8	M	**8**	10	M
9	8	M	**9**	14	F
10	9	M	**10**	15	M
11	9	M	**11**	10	M
12	8	M	**12**	6	M
13	6	M	**13**	12	M
14	12	M	**14**	8	M
15	8	M	**15**	12	M

8.3 Results

8.3.1 H1: Using the selected telepractice system, the telepractice stuttering assessment will be feasible and comparable to the face-to-face assessment.

A total of 30 assessment sessions were carried out; fifteen face-to-face sessions and a further 15 TP sessions. All 30 sessions were successfully carried out, and no session was terminated due to technical issues or any other difficulty. However, there were instances when difficulties occurred, e.g. the distant clinician losing sight of the child when observing a face-to-face session as the child changed his position. This also occurred once in a face-to-face-led session. In another session, a child spoke in an extremely quiet voice as a form of stutter avoidance, making understanding him difficult for both the distant and face-to-face SLTs (this occurred in a TP-led session). In terms of technical issues, four sessions were severely interrupted. However, to overcome this, the connection was re-established, which enhanced the quality of the connection, although not to an optimal level. As for audio and visual quality ratings, audio quality was rated as excellent in 80% of the sessions (24 sessions), very good 6.7% of the time (2 sessions), good 3.3% of the time (1 session) and fair 10% of the time (3 sessions), with no ratings of poor. In this sense, audio difficulties related to delays in signal reception and interruption of audio. As for visual quality, 56% of the sessions were rated excellent (17 sessions), 16.7% were rated very good (5 sessions), 13.3% were rated good (4 sessions), fair 3.3% of the time (1 session), and 10% were rated poor (3 sessions) (see Appendix 10). However, such instances of technical issues did not preclude carrying out the sessions. Most of the technical issues were related to visual quality, whereas audio was stable in most of the sessions.

8.3.2 H2: Results from the assessment will present a high degree of clinically acceptable levels of agreement and reliability between the two environments, validating the telepractice modality for stuttering assessment.

The outcome measures used for assessment were:

1- Percentage of syllables stuttered (%SS);
2- SSI-IV: %SS subcomponent, the duration subcomponent, the physical concomitants subcomponent, the total score, and the final severity ratings; and
3- Stuttering Severity Rating (SR) scale.

To answer this question, several statistical analyses were utilised. The Wilcoxon signed rank test of difference was used to compare scores in the face-to-face and TP environments. Further, the Bland and Altman method (Bland & Altman, 2003) was used to assess the level of agreement of findings retrieved from the face-to-face and TP sessions, except for the final severity ranking, where quadric weighted kappa (k_W) was used. In addition, percentage of exact agreement, and percentage agreement within 1, 2, and 3 sub-score values of telepractice and face-to-face ratings of SSI-IV subcomponents (%SS subcomponent, physical concomitants subcomponent, and total score) were obtained, with the exception of the duration subcomponent of the SSI-IV, where only exact percentage agreement and agreement within 2 sub-scores were obtained, and percentage agreement within 1 sub-score value was obtained for SR, so as to be able to compare the results with other studies. Finally, the reliability of telepractice and face-to-face data was investigated using the ICC (2,1) for all measures, except for the final severity rating of the SSI-IV, where k_W was employed.

8.3.2.1 Percentage of syllables stuttered (%SS):

A Wilcoxon signed rank test found significant differences in the scores for face-to-face (median=6.85) and telepractice (median=6.80) conditions, Z=-2.09, p=.04 (see Figure 5).

The Bland and Altman method (see Table 12 and Figure 7) revealed that the mean difference was .82% higher in face-to-face ratings when compared to TP ratings, which is considered as a clinically insignificant difference. However, the LOAs were wide, ranging from -4.33% to 5.97%, which implies that TP ratings of %SS could potentially be 4.33% less or 5.97% more than face-to-face ratings.

After inspection of the plot, the data points with the largest differences were ratings of four participants that had internet connection problems; therefore, it was decided to re-analyse this dataset without these four data points. To be clear, only data with large differences due to technical difficulties were excluded, whereas those that returned large differences but did not result from any technical difficulty were retained. We believe that analysing data before and after exclusion is of value, as it is a form of sensitivity analysis.

Analysis of %SS after excluding samples

A Wilcoxon signed rank test found no significant difference in the scores for face-to-face (median=6.65) and telepractice (median=6.80) conditions, Z=-0.92 , p=0.36 (see Figure 6).

Under these conditions, LOA of %SS vastly improved, as demonstrated by the Bland and Altman plot (see Figure 8), where LOAs were between -2.72% and 3.04%. This implies that ratings of %SS via the TP model could potentially be 2.72% less or 3.04% more than face-to-face ratings. The mean value of the face-to-face scoring of %SS is 0.16% higher when compared to online ratings.

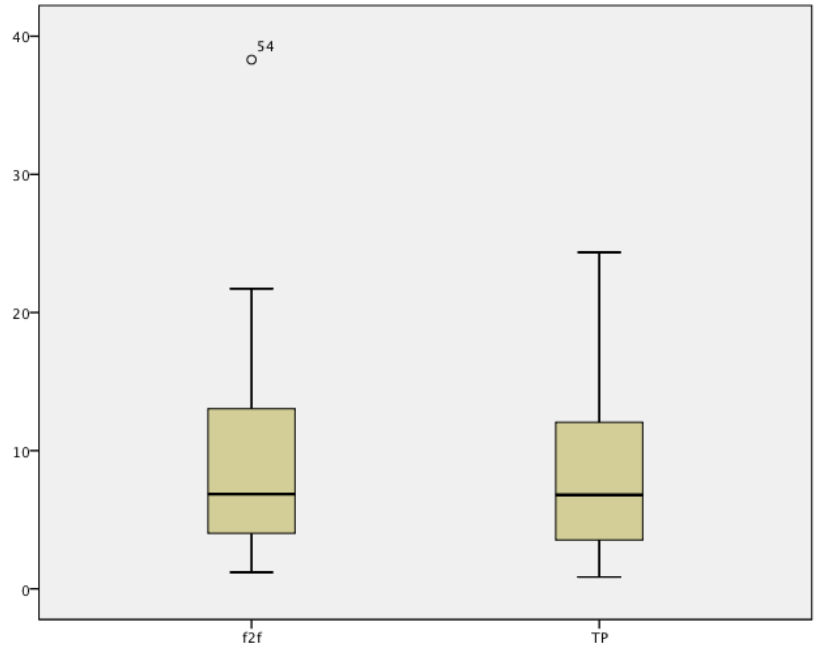


Figure 5 Box-Plots of the Scores of %SS in Face-to-Face and TP Environments before Conducting Sensitivity Analysis

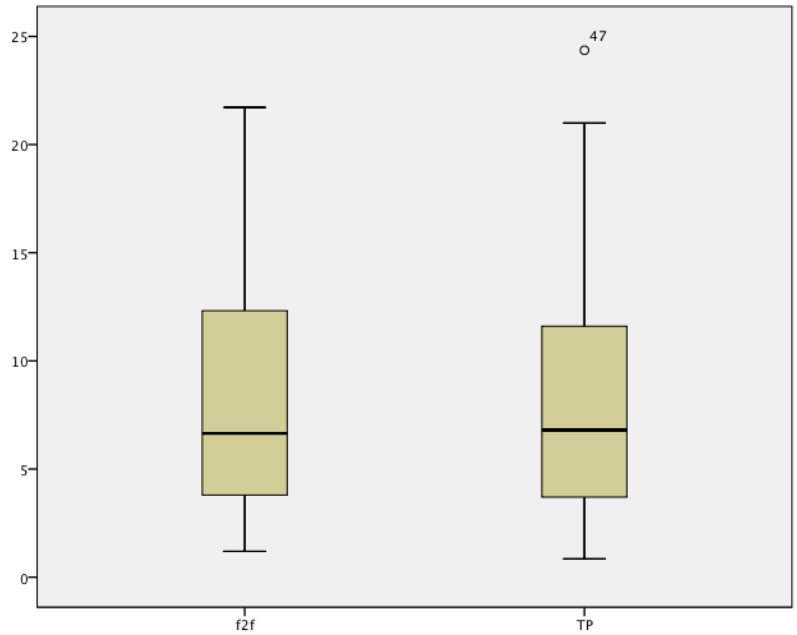


Figure 6 Box-Plots of the Scores of %SS in Face-to-Face and TP Environments after Conducting Sensitivity Analysis

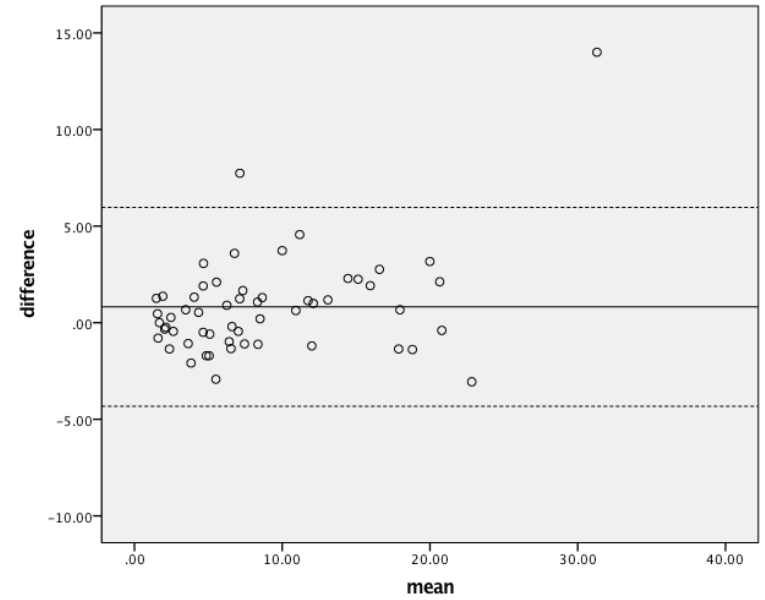


Figure 7 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of %SS before Conducting Sensitivity Analysis. The Two Dotted Lines Define the LOAs (Mean of Difference ± 1.96 SD).

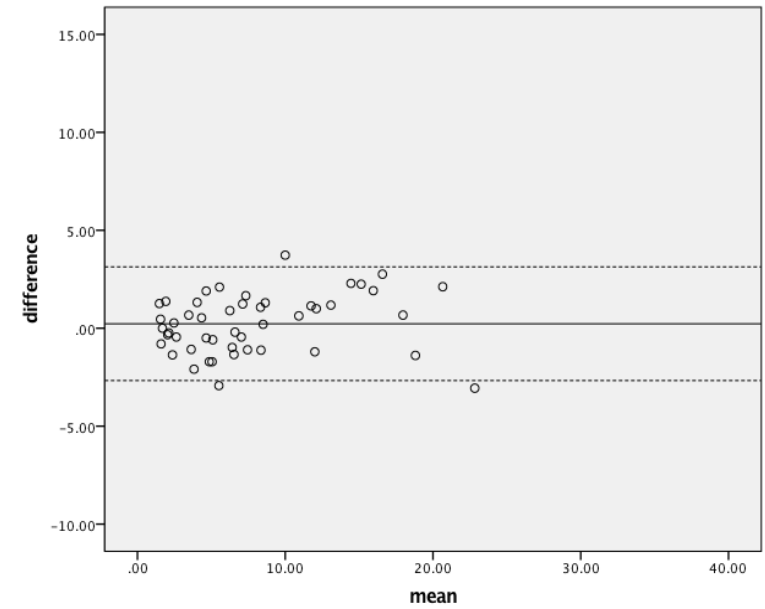


Figure 8 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of %SS after Conducting Sensitivity Analysis. The Two Dotted Lines Define the LOAs (Mean of Difference ± 1.96 SD).

8.3.2.2 Stuttering Severity Instrument (SSI-IV)

SSI-IV %SS subcomponent

A Wilcoxon signed rank test found no significant difference in the scores for face-to-face (median=12.5) and telepractice (median= 12.5) conditions, Z=-1.18 , p=.24 (see Figure 9).

The Bland and Altman method (see Table 13) revealed that the mean difference was .35 higher in face-to-face ratings when compared to TP ratings, which is considered a clinically insignificant difference. On the other hand, LOAs were wide, ranging from -2.3 to 3, which implies that, using the TP ratings of %SS subcomponent, LOAs could potentially be up to 2.3 points less or 3 points more than face-to-face ratings (see Figure 10). However, the percentage agreement for exact score was 46.6%, while it was 70%, 93%, and 100% for 1, 2, and 3 points difference respectively, which is comparable to the agreement estimates of the SSI-IV calculated in Chapter 5.

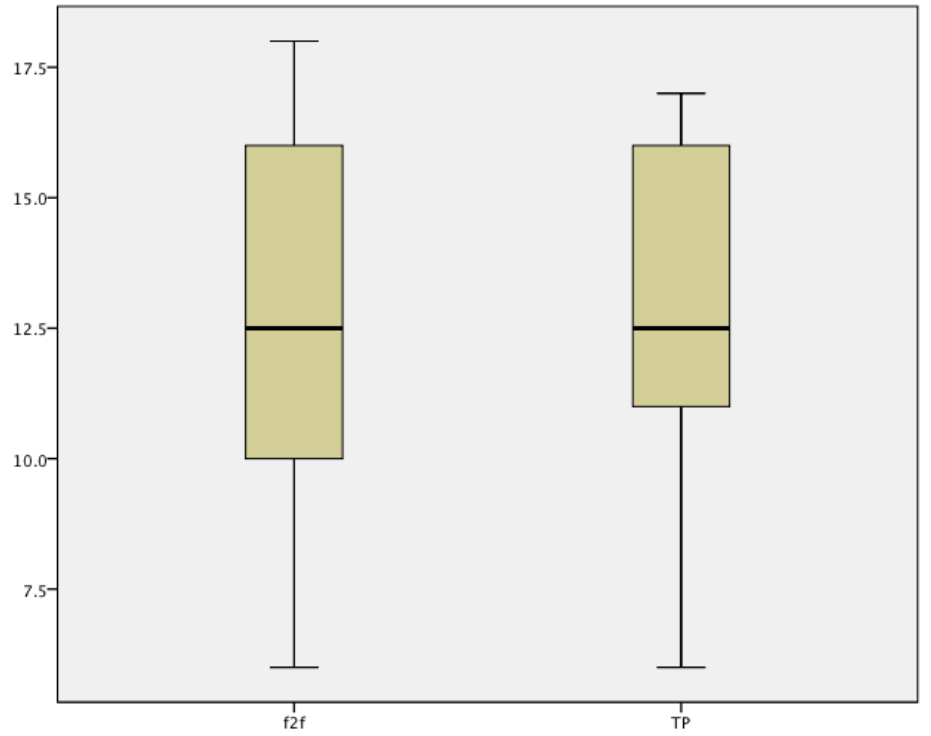


Figure 9 Box-Plots of the Scores of %SS Subcomponent in Face-to-Face and TP Environments.

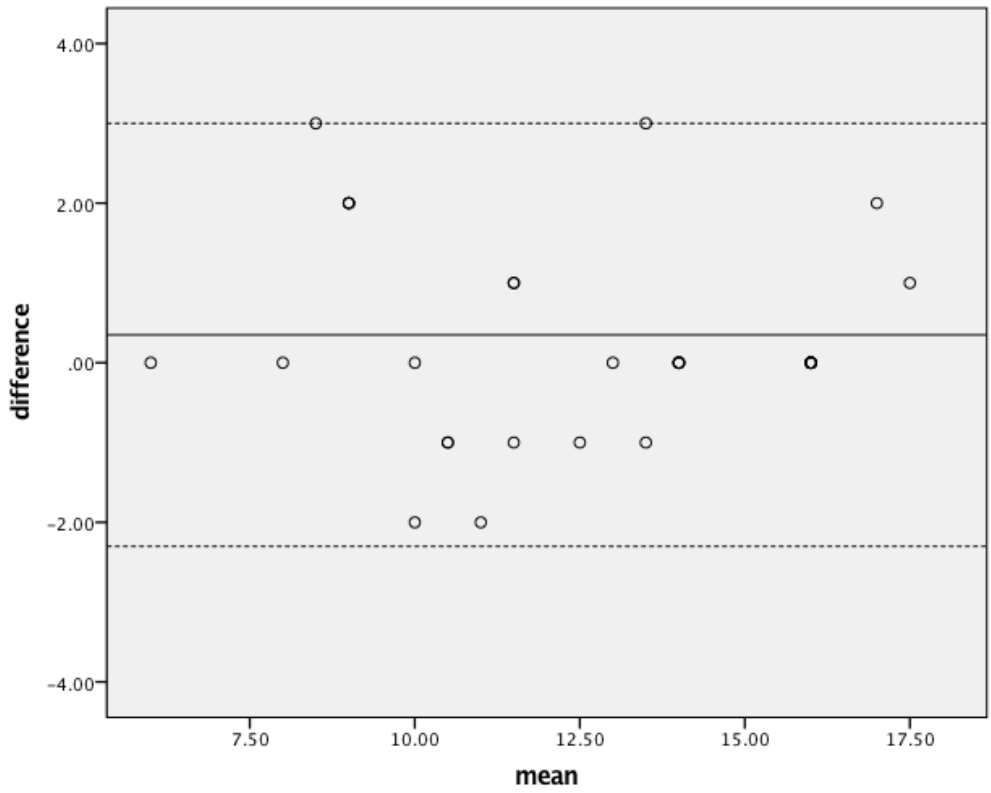


Figure 10 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of the %SS Sub-Score. The Central Line Represents the Mean Differences between Face-to-Face and TP Ratings. The Two Dotted Lines Define the LOAs (Mean of Difference ± 1.96 SD)

SSI-IV duration subcomponent

A Wilcoxon signed rank test found no significant difference in the scores for face-to-face (median=6) and telepractice (median=8) conditions, Z= --1.15, p=.24 (see Figure 11).

The Bland and Altman method (see Table 13) revealed that the mean difference was 0.27 points lower in face-to-face ratings when compared to TP ratings, which is considered a clinically insignificant difference. On the other hand, LOAs were wide, ranging from -2.73 to 2.19, which implies that, using the TP ratings, the duration subcomponent could potentially be up to 2.73 points less or 2.19 points more than face-to-face ratings (see Figure 12). However, the percentage agreement of exact score was 60% while, for within 2 sub-score values, it was 100%, which is comparable to the agreement estimates of the SSI-IV calculated in Chapter 5.

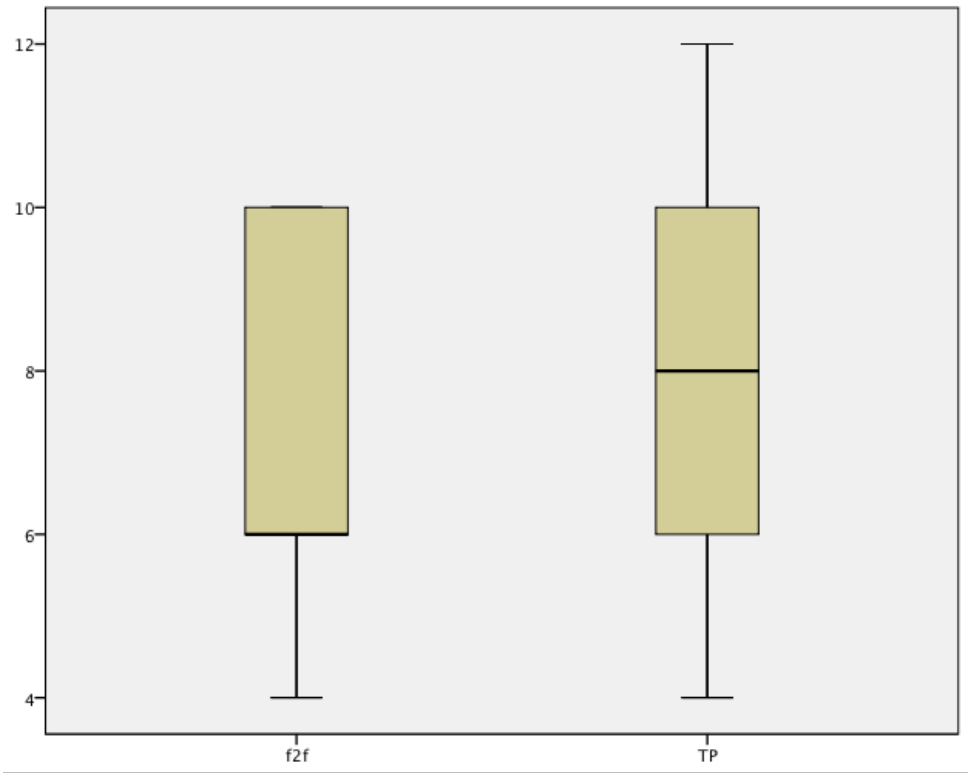


Figure 11 Box-Plots of the Scores of the Duration Subcomponent in Face-to-Face and TP Environments.

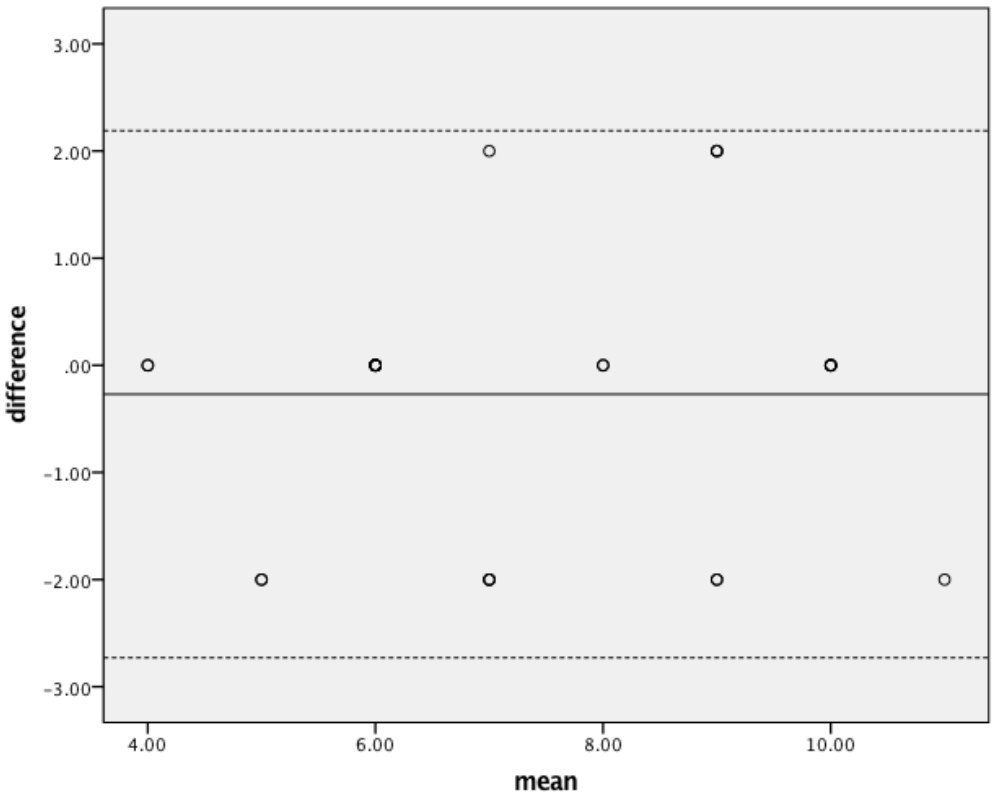


Figure 12 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of the Duration Subcomponent. The Central Line Represents the Mean Differences between Face-to-Face and TP Ratings. The Two Dotted Lines Define the LOAs (Mean of Difference ± 1.96 SD)

Physical concomitants subcomponent

The Wilcoxon signed rank test found no significant difference in the scores for face-to-face (median=2) and telepractice (median=1.5) conditions, Z= -1.8, p= 0.07 (see Figure 13).

The Bland and Altman method (see Table 13) revealed that the mean difference is .60 points higher in face-to-face ratings when compared to TP ratings, which is considered a clinically insignificant difference. Despite this, LOAs were wide, ranging from -2.9 to 4.1, which implies that TP ratings of physical concomitants could potentially be 2.9 points less or 4.1 points more than face-to-face ratings (see Figure 14). However, percentage agreement of exact score was 26.6%, whereas it was 53.3%, 86.6%, and 96.6% for within 1, 2, and 3 sub-score values respectively, which is comparable to the agreement estimates of the SSI-IV calculated in Chapter 5.

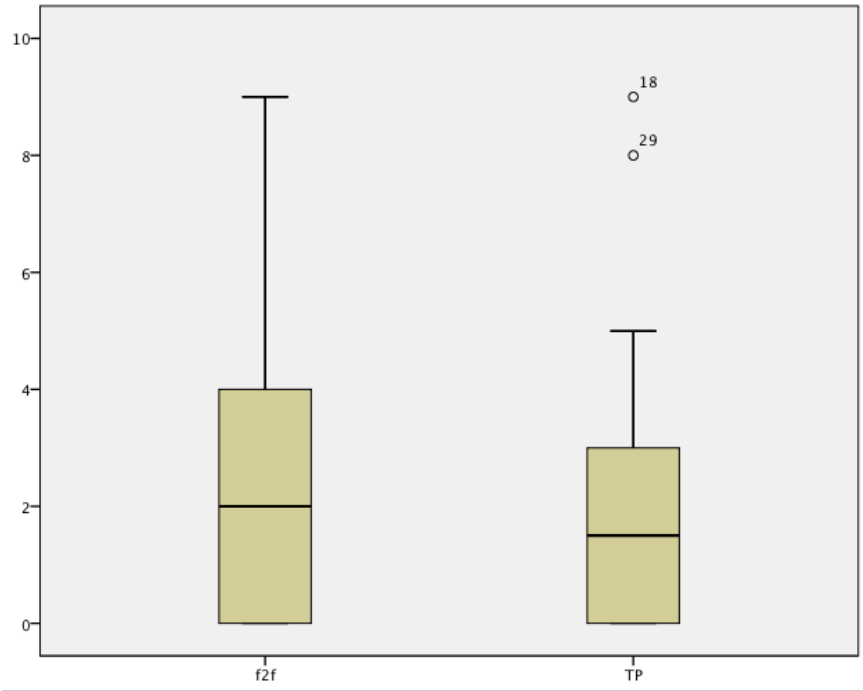


Figure 13 Box-Plots of the Scores of Physical Concomitants in Face-to-Face and TP Environments.

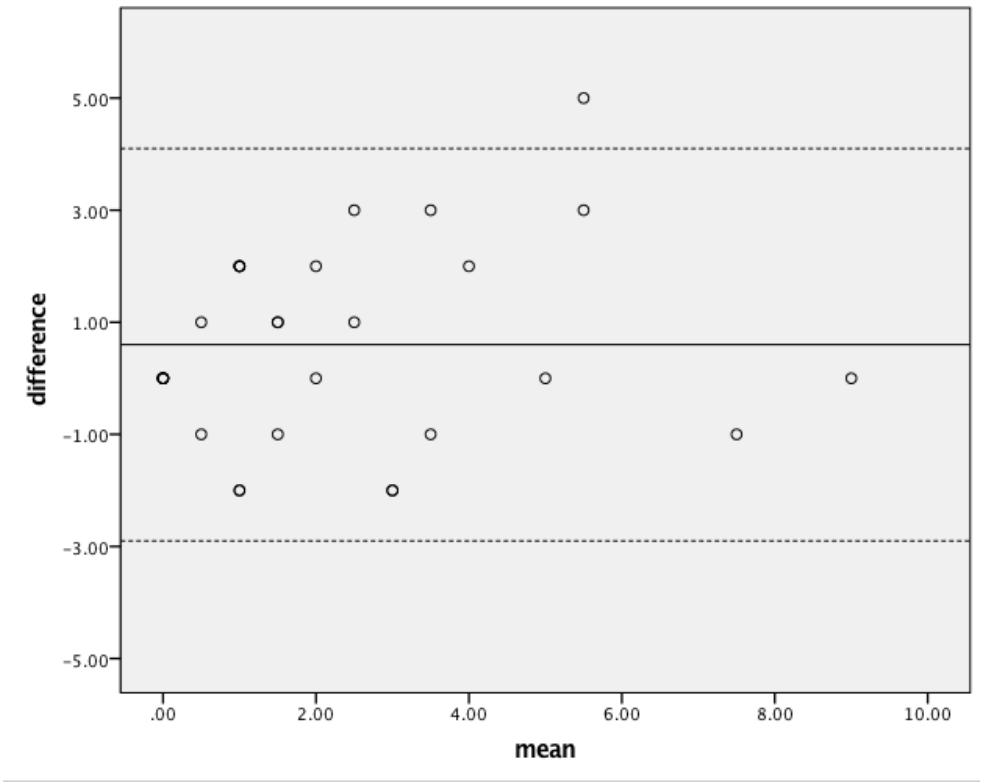


Figure 14 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of Physical Concomitants. The Central Line Represents the Mean Differences between Face-to-Face and TP Ratings. The Two Dotted Lines Define the LOAs (Mean of Difference ± 1.96 SD)

SSI-IV total score

The Wilcoxon signed rank test found no significant difference in the scores for face-to-face (median= 21) and telepractice (median= 20) conditions, Z= -1.06, p= 0.28 (see Figure 15).

The Bland and Altman method (see Table 13) revealed that the mean difference is .60 points higher in face-to-face ratings when compared to TP ratings, which is considered a clinically insignificant difference. Despite this, LOAs were wide, ranging from -4.71 to 5.9, which implies that TP ratings of total score could potentially be 4.71 points less or 5.9 points more than face-to-face ratings (see Figure 16). However, percentage agreement of exact score was 26.6%, whereas it was 50%, 66.6%, and 80% for within 1, 2, and 3 sub-score values respectively, which is comparable to the agreement estimates of the SSI-IV calculated in Chapter 5.

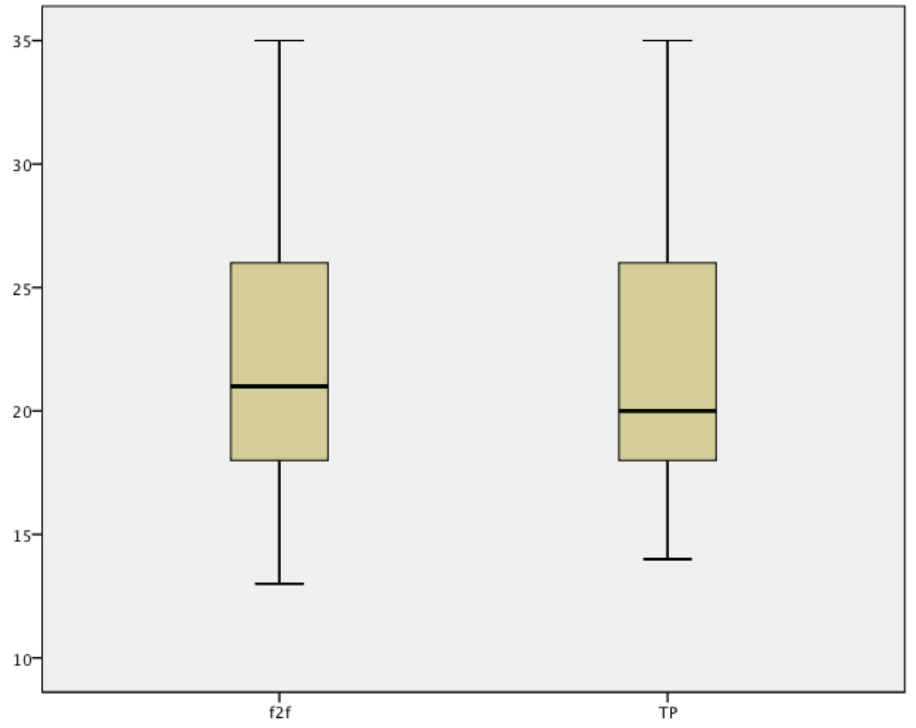


Figure 15 Box-Plots of the SSI-IV Total Scores in Face-to-Face and TP Environments

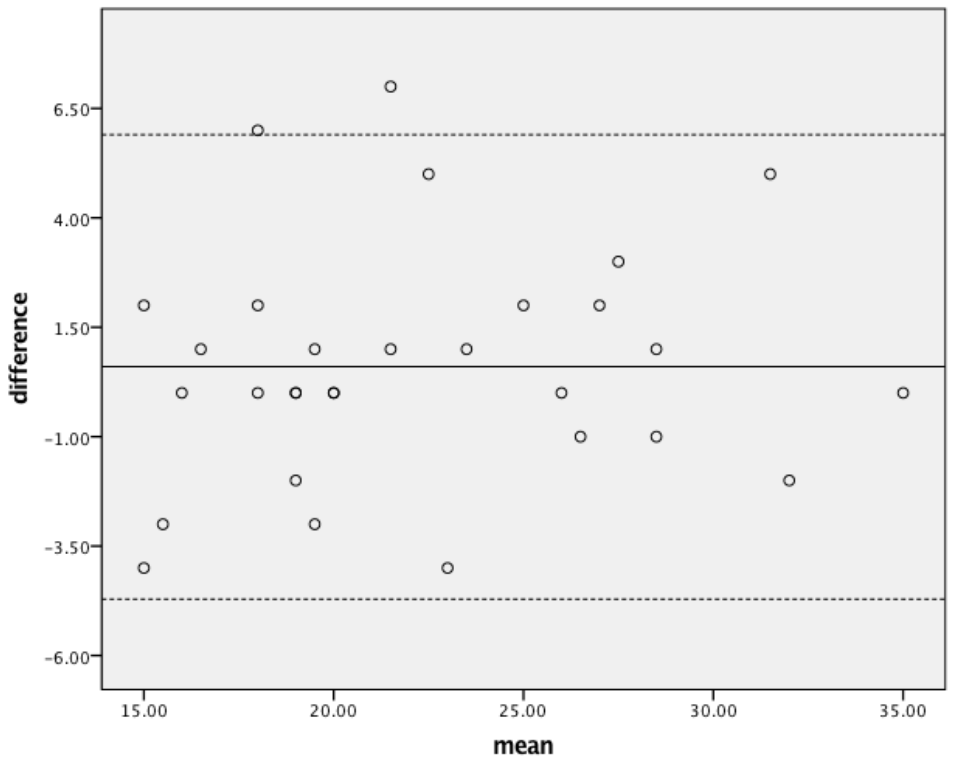


Figure 16 Bland–Altman Plot of the Difference in TP and Face-to-Face Scores against the Mean Scores of Total Scores. The Central Line Represents the Mean Differences between Face-to-Face and TP Ratings. The Two Dotted Lines Define the LOAs (Mean of Difference ± 1.96 SD)

SSI-IV final severity rating

Quadric weighted kappa statistics indicated very good agreement between face-to-face and TP final severity ratings ($k_{W=}$ .81, 95% CI [0.69-0.96]).

8.3.2.3 SR

The Wilcoxon signed rank test found no significant difference in the scores for face-to-face (median=5) and telepractice (median= 4) conditions, Z= -0.64, p= 0.51 (see Figure 17).

The Bland and Altman method (see Table 12) revealed that the mean difference is .16 points higher in face-to-face ratings when compared to TP ratings, which is considered a clinically insignificant difference. Despite this, LOAs were wide, ranging from -2.54 to 2.9, which implies that TP ratings of SR could potentially be 2.54 points less or 2.9 points more than face-to-face ratings (see Figure 18). However, 87.7% of ratings were within the accepted 1 scale value of difference (Eve et al., 1995; O'Brian et al., 2004; Onslow, Andrews & Costa, 1990).

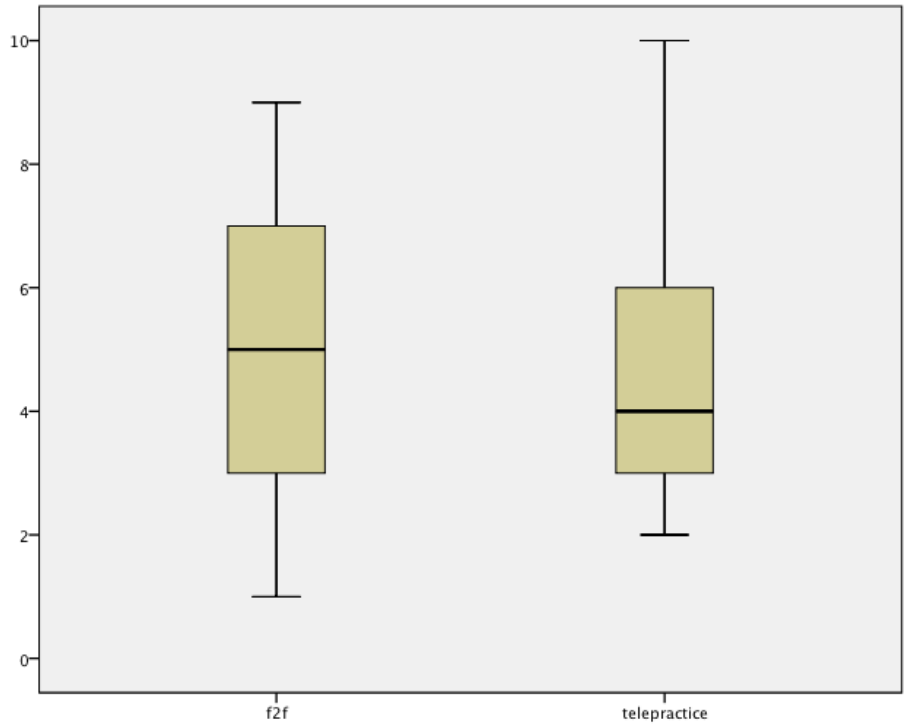


Figure 17 Box-Plots of SR in Face-to-Face and TP Environments

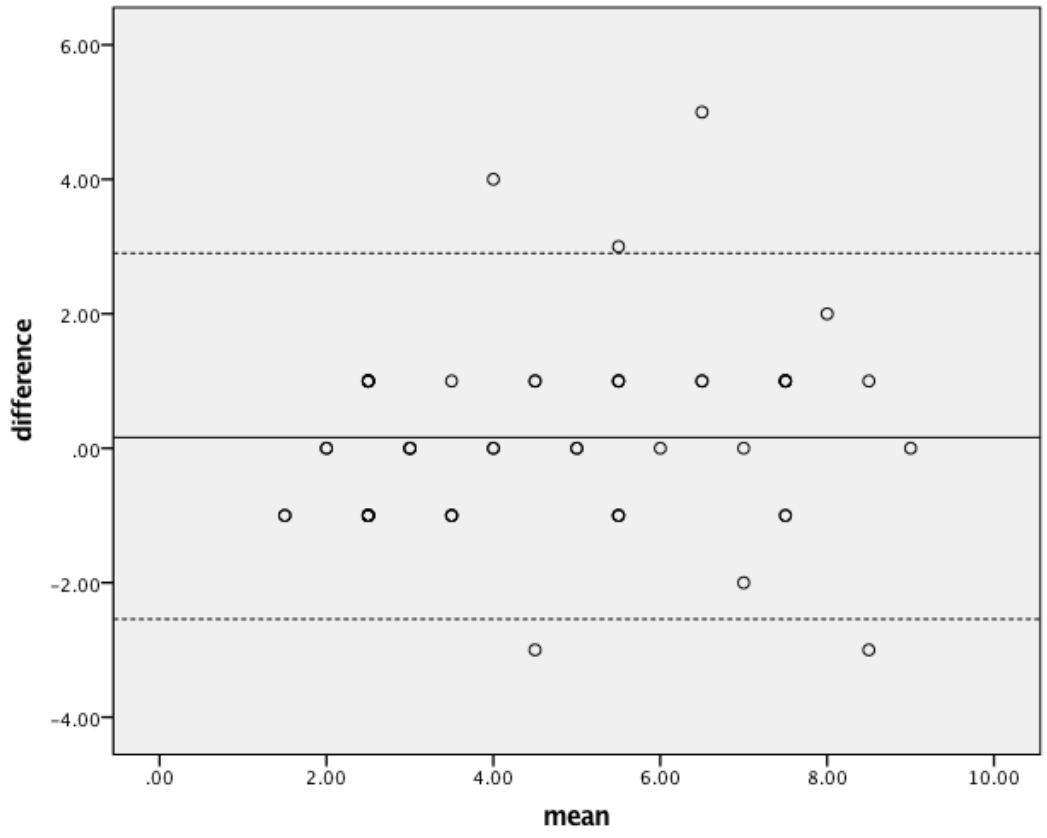


Figure 18 Bland–Altman Plot of the Difference in TP and Face-to-Face SR Scores against the Mean Scores of SR. The Central Line Represents the Mean Differences between Face-to-Face and TP Ratings. The Two Dotted Lines Define the LOAs (Mean of Difference ± 1.96 SD)

8.3.2.4 Reliability of data

Reliability of the face-to-face data (see Table 14)

Reliability of %SS face-to-face ratings

Intra-rater reliability:

ICC indicated very high intra-rater reliability of %SS (ICC=0.99).

Inter-rater reliability:

ICC indicated very high inter-rater reliability of %SS (ICC=0.92).

The reliability of SSI-IV face-to-face ratings

Please refer to Chapter 5

The reliability of SR face-to-face ratings

Intra-rater reliability:

ICC indicated very high intra-rater reliability of SR (ICC= .94).

Inter-rater reliability:

ICC indicated high inter-rater reliability of SR (ICC= .88).

Reliability of the TP system data (see Tables 15 and 16)

The reliability of %SS TP ratings

Intra-rater reliability:

ICC indicated very high intra-rater reliability of %SS (ICC=.97).

Inter-rater reliability:

ICC indicated very high inter-rater reliability of %SS (ICC=.94).

The reliability of SSI-IV TP ratings

Intra-rater reliability:

ICC ranged between .91 and .97, suggesting very high intra-rater reliability of all SSI subcomponents.

The intra-rater reliability for final severity rating was (k_W=.68, 95% CI .49 to .87), suggesting good reliability.

Inter-rater reliability:

Inter-rater reliability ranged from moderate to very high (ICC=.67 to .95) for all SSI components; the least favourable ICC value was for the physical concomitants subsection (ICC=.67).

The inter-rater reliability for severity rating was (k_W=.69, 95% CI .34 - 1), suggesting good reliability.

The reliability of SR TP ratings

Intra-rater reliability:

ICC indicated high intra-rater reliability of SR (ICC= .86).

Inter-rater reliability:

ICC indicated very high inter-rater reliability of SR (ICC= .82).

Table 12 LOAs of %SS and SR

Assessment Task	Mean Difference	SD of Difference	95% CI of the Mean Difference	LOA
%SS *	.82	2.63	.11 to 1.5	-4.33 to 5.97
SR	.16	1.42	-.22 to .54	-2.54 to 2.9

*The LOA of %SS reported in the table is before conducting sensitivity analysis. After sensitivity analysis, LOA ranged from -2.27 to 3.04, with a mean difference of .16 (95% CI= -0.25 to 0.57) and SD of difference=1.47.

Table 13 LOAs of SSI-IV Subcomponents

SSI-IV Subcomponent	Mean Difference	SD of Difference	95% CI of the Mean of Difference	LOA
%SS	.35	1.35	-1.2 to 0.82	-2.3 to 3.0
Duration	-.27	1.26	-.73 to 0.20	-2.73 to 2.19
Physical Concomitants	.60	1.80	-.05 to 1.25	-2.90 to 4.1
Total score	.60	2.71	-.41 to 1.61	-4.71 to 5.91

Table 14 Inter-/Intra-Rater Reliability of Face-to-Face Ratings of the %SS and SR: ICC (95% CI).

Assessment Task	Face-to-Face Intra-Rater Reliability	Face-to-Face Inter-Rater Reliability
%SS	.99 (0.98-0.99)	.92 (.67-.97)
SR	.94 (.87-.98)	.88 (.72-.95)

Table 15 Inter-/Intra-Rater Reliability of TP Ratings of the %SS and SR: ICC (95% CI)

Assessment Task	TP Intra-Rater Reliability	TP Inter-Rater Reliability
%SS	.97 (.94 .99)	.94 (.85-.97)
SR	.86 (.68-.94)	.82 (.60-.93)

Table 16 Inter-/Intra Rater Reliability of TP Ratings of the SSI-IV Subcomponents: ICC (95% CI)

SSI-IV Subcomponent	TP Intra-Rater Reliability	TP Inter-Rater Reliability
%SS	.93 (.72-.98)	.85 (.52-.96)
Duration	.89 (.64-.97)	.75 (.47-.93)
Physical Concomitants	.95 (.84-.98)	.67 (.21-.90)
Total score	.94 (.76-.98)	.95 (.81-.98)

8.3.3 H3: Children and their parents will feel positive about the potential of using telepractice for stuttering assessment, satisfied with the telepractice assessment of their stuttering, and willing to use the telepractice system for stuttering therapy in the future.

Parents and children completed two questionnaires containing qualitative and quantitative information (in the form of a 5-point Likert scale measurement): a pre-assessment questionnaire focused on respondents' preconceptions about assessment via telepractice, and a second returned feedback form from both parents and children about their satisfaction with the process once the child had undergone the TP assessment (See Appendix 8).

For ease of interpretation, before conducting statistical analysis, parents' and children's responses across the 5-point Likert scale were collapsed to form three groups: 'strongly disagree' was grouped with 'disagree' in a category labelled 'disagree'; 'I don't know/neutral' was kept the same; and 'agree' and 'strongly agree' were grouped in a category labelled 'agree'. Moreover, patterns of responses across the pre-and post-assessments of both parents and children were reported descriptively. Finally, to assess the extent of change in perceptions between the pre- and post-assessment, a Wilcoxon matched pairs signed rank test was applied. All parents (15 respondents, 12 mothers, and 3 fathers) and children (14 respondents) in the sample indicated that they had never used a TP model for any medical service.

8.3.3.1 Parental Responses

Pre-Assessment

Prior to the assessment, most parents reported their willingness to use telepractice if it was an available option, and they were comfortable with their children undergoing an online assessment (93.3%), except for one parent, who was not sure. Moreover, 86.6% of parents reported that they would consider the TP model for stuttering treatment, and 73.3% believed that their children would enjoy the assessment task online. In terms of audio and video quality, most parents were confident that they would not experience any difficulty in seeing (80%) or hearing (86.7%) the distance therapist, with the rest being unsure. In addition, 73.3% agreed that instructions would be easy for their children to follow and their children would be given enough time to complete tasks, with the rest being unsure (26.7%). In addition, 86.7% agreed that the TP model would allow easy access to therapy services for

CWS, in addition to also agreeing that the model would save time and money. However, when parents were asked whether they rated online assessment as equivalent to a face-to-face assessment, opinions were divided: 46.7% were not sure, 46.6% agreed, and only one parent (6.7%) disagreed. This variability in responses was observed in another question where parents were asked if they would prefer the traditional method of face-to-face assessment over the TP model, wherein 46.6% were unsure, 13.3% indicated their preference for the traditional method over the TP model, and 40% disagreed about the superiority of the traditional method over the TP model.

Post-Assessment

After running the statistical analysis, some changes in responses were observed after conducting the assessment, with the majority remaining the same (see Table 17). The Wilcoxon test (Questions 1, 2, and 3) revealed that there was no significant change in the response to questions addressing levels of comfort using the TP model, as similarly high levels of comfort remained post-assessment. There was also no significant change in perceptions regarding audio and video quality (Questions 5 and 6), with perceptions being positive both pre- and post-assessment. In Question 4, the Wilcoxon test revealed a significant change. Parents were asked if they thought their children enjoyed the session, and affirmative responses increased from 73.3 to 100% post-assessment. Additionally, when asked whether the parents rated the online assessment equivalent to face-to-face assessments (Question 7), unlike the pre-assessment, when opinions were more divided, agreement significantly increased from 46.6 to 93.3%. When asked about the clarity of instructions and whether their children were given enough time to respond (Questions 8 and 9), the Wilcoxon test revealed that parent's agreement significant increased from 73.3 to 100%. Although not statically significant, parents' preference for telepractice over face-to-face practice increased from 40 to 73.3% (Question 13). Finally, the Wilcoxon test revealed no significant change in Questions 11–12, wherein most parents believed that the TP model saves them time and money and allows easy access to SLT services for CWS pre- and post-assessment.

Additional Parents' Comments

Additional comments from 15 parents who had undergone a TP-led assessment are summarised below. The complete range of comments can be seen in Appendix 11. The most common comments related to convenience; in this sense, three parents stated that the TP model was convenient. Four parents said that the TP model required less travel and was time

saving. Two parents mentioned the word 'flexible', one said that it saved travel time, and another parent mentioned that it did not require his physical presence with his son. Moreover, parents commented that the model was engaging, e.g., 'My daughter engaged perfectly with the distant clinician' (PCH5); 'I have never seen my child that focused and engaged with the clinician … compared to face-to-face sessions' (PCH11). Three parents remarked that this model might be less stressful for their children, as the children would be seen in their own environments. Three parents replied that they were happy with the TP model, as it follows technological advances. Two parents commented that the TP model offers more access to therapy: 'This model can help services reach as many people as possible' (PCH5); 'This model allows services to reach people who can't travel' (PCH1). In addition, PCH1 said the model could offer more privacy: 'I think this model allows more privacy and makes people talk, as they will be communicating from a relaxed environment, from their homes'. Fewer remarks were received regarding drawbacks of the TP model. One parent talked about establishing rapport, preferring her child to be seen face-to-face initially and maybe seen through telepractice afterwards. Two parents stated that the model may be affected by slow internet speeds.

Children

14 children, with an age range of 8–15 years (M=10.8; SD=2.1), filled in the questionnaires before and after the TP assessments.

Pre-Assessment

Prior to the assessment, most children (85.7%) indicated that they were comfortable undergoing stuttering assessments online and that they would be willing to use it if available. Moreover, 92.2% agreed that they expected to enjoy assessment tasks carried out online. However, when asked if they might consider the TP model for treatment, opinions were more divergent, as 42.9% were not sure and 57.1% agreed. When children were asked about their expectations regarding visual and audio quality, all children agreed that they would not have any problem hearing the clinician during an online session. When asked about visual quality, 78.6% agreed, one child was not sure, and two children disagreed. Of all the respondents, 92.2% of the children were confident that instructions during the TP session would be clear and easy to follow. They believed that the TP model enabled easy access to assessment and therapy for CWS. Although most children agreed that instructions would be clear and easy to follow, fewer children agreed that they would be given sufficient time to follow instructions

through this model (64.3%), with the remainder being unsure (35.7%). Additionally, 64.3% of the children agreed that assessment through the TP model is a satisfactory alternative to face-to-face practice, whereas the rest were unsure (28.6%) or disagreed (7.1%). Finally, when children were asked if they preferred the traditional assessment model over telepractice, 50% disagreed, 14.3% were unsure, and 35.7% agreed that they preferred the traditional model.

Post-Assessment

After running the statistical analysis, some changes in responses were observed after the assessment (see Table 18), with the majority remaining the same, particularly regarding Questions 3, 7, 9, and 12. In Question 3, when asked if they were happy during the online session and whether telepractice would be considered for therapy, significantly higher acceptance was observed compared to the pre-assessment, with agreement increasing from 57.1 to 92.2%. After carrying out the assessment, there was a significant change in children's opinions regarding whether they thought TP assessment was equivalent to face-to-face assessment (Question 7), with 92.9% agreeing, compared to 50% from the previous assessment. In addition, more children agreed to Question 9 (92.2%), where they were asked if they had sufficient time to respond to and follow instructions. Finally, for Question 12, when asked if they preferred the traditional face-to-face model over the TP model, unlike in the pre-assessment where opinions diverged, the majority disagreed (92.2%). Generally, changes in perceptions of both parents and children, statistically significant or not, were all positive, indicating acceptance of the TP model (see Appendix 2).

Additional Children's Comments

Additional comments from 14 children who had undergone a TP-led assessment are summarised below. The complete range of comments can be seen in Appendix 12. The most frequent comment was that the TP model was 'less stressful' than the face-to-face model, as the therapist was not physically present in the room. Moreover, with the exception of one child, who commented that he felt nervous during the TP session, children reported that it is easier for them to communicate through this model: 'I preferred the telepractice method over the face-to-face because it felt less stressful seeing the therapist in a different location and through the screen. My hands usually sweat from stress when I talk to my SLT' (CH3); 'This model was less stressful, and it was easier to communicate with the clinician. I can express myself more with this model' (CH8).

Children also discussed the convenience of the TP model and mentioned that it was time saving and did not require travelling: ‘I came from outside Riyadh to attend the appointment. This option will save me time. Although I do not have internet at my house, I can go to my grandparents’ house or the closest hospital to my house’ (CH1); ‘It is very convenient, as it saves travelling time. I get the session within my home environment. I can avoid traffic’ (CH7).

Table 17 Results of Pre- and Post-Assessment Questionnaires of Parents Condensed from a 5-Point Likert Scale to a 3-Point Likert Scale to Reveal Basic Groups of 'Disagree', 'I Don't Know, and 'Agree'. The Italics and Brackets Indicate Pre-/Post-Assessment Conditions

Question	Pre-Assessment			Post –Assessment			Z	P Value
	Disagree	Don't Know	Agree	Disagree	Don't know	Agree		
1- I *will be* (I am) happy to use telepractice if it is available in the hospital or healthcare facility nearest to my place of residence.	0	1 (6.7%)	14 (93.3%)	0	0	15 (100%)	-1.00	.317
2- I *am* (was) comfortable for my child to undergo stuttering assessment via the internet.	0	1 (6.7%)	14 (93.3%)	0	0	15 (100%)	-1.00	.317
3- I *will be* (was) happy for my child to be online and would consider using the internet for stuttering treatment.	0	2 (13.3%)	13 (86.6%)	1 (6.7%)	1 (6.7%)	13 (86.7%)	-.447	.655
4- I believe my child *will enjoy* (enjoyed) carrying out assessment tasks online.	0	4 (26.7%)	11 (73.3%)	0	0	15 (100%)	-2.00	.046*
5- *I will have* (had) no difficulty in seeing the online speech therapist via the telepractice video link.	0	3 (20%)	12 (80%)	1 (6.7%)	0	14 (93.3%)	-.378	.705
6- *I will have* (had) no difficulty hearing the online speech therapist via the telepractice video link.	0	2 (13.3%)	13 (86.7%)	1 (6.7%)	0	14 (93.3%)	0	1.00
7- I would rate the online assessment as being equal to an assessment conducted traditionally in the face-to-face method.	1 (6.7%)	7 (46.7%)	7 (46.7%)	1 (6.7%)	0	14 (93.3%)	-2.646	.008*
8- The instructions given during the online assessment *will be* (were) clear and easy to follow.	0	4 (26.7%)	11 (73.3%)	0	0	15 (100%)	-2.00	.046*
9- . My child *will have* (had) sufficient time to follow the instructions given during the assessment.	0	4 (26.7%)	11 (73.3%)	0	0	15 (100%)	-2.00	.046*
10- Telepractice can satisfactorily replace a face-to-face assessment.	2 (13.3%)	3 (20%)	10 (66.7%)	1 (6.7%)	1 (6.7%)	13 (86.7%)	-1.633	.102
11- Telepractice *will allow* (can allow) easy access to assessment and therapy for children who stutter.	0	2 (13.3%)	13 (86.7%)	0	1 (6.7%)	14 (93.3%)	-1.00	.317
12- Telepractice will save me travelling time and money.	0	2 (13.3%)	13 (86.7%)	0	0	15 (100%)	-1.414	.157
13- I would prefer for my child to have a traditional (face-to-face) assessment with the speech therapist.	6 (40%)	7 (46.7%)	2 (13.3%)	11 (73.3%)	2 (13.3%)	2 (13.3%)	-1.318	.187

Table 18 Results of Pre- and Post-Assessment Questionnaires of Children Condensed from a 5-Point Likert Scale to a 3-Point Likert Scale to Reveal Basic Groups of 'Disagree', 'I Don't Know', and 'Agree'. The Italics and Brackets Indicate Pre-/Post-Assessment Conditions

Question	Pre-Assessment			Post-Assessment			Z	P value
	Disagree	I Don't Know	Agree	Disagree	I Don't Know	Agree		
1-I *will be* (I am) happy to use telepractice if it is available in the hospital or healthcare facility nearest to my place of residence.	0	2 (14.3%)	12 (85.7%)	0	0	14 (100%)	-1.414	.157
2-I *am* (was) comfortable to undergo a stuttering assessment via the internet.	0	2 (14.3%)	12 (85.7%)	0	0	14 (100%)	-1.414	.157
3-I anticipate I *will be* (was) happy being online and would consider using the internet for stuttering treatment.	0	6 (42.9%)	8 (57.1%)	0	1 (7.1%)	13 (92.2%)	-2.236	.025*
4-*I anticipate I will enjoy* (I enjoyed) carrying out assessment tasks online.	0	1 (7.1%)	13 (92.2%)	0	0	14 (100%)	-1.000	.317
5- I anticipate I *will have* (I had) no difficulty in seeing the online speech therapist via the telepractice video link.	2 (14.3%)	1 (7.1%)	11 (78.6%)	4 (28.6%)	0	10 (71.4%)	-.849	.369
6- I *will have* (had) no difficulty hearing the online speech therapist via the telepractice video link.	0	0	14 (100%)	0	0	14 (100%)	.000	1.00
7-I would rate the online assessment as being equal to an assessment conducted traditionally in the face-to-face method.	1 (7.1%)	6 (42.9%)	7 (50%)	0	1 (7.1%)	13 (92.9%)	-2.111	.035*
8-The instructions given during the online assessment *will be* (were) clear and easy to follow	0	1 (7.1%)	13 (92.2%)	0	2 (14.3%)	12 (85.7%)	-.577	.564
9- I *expect to have* (had) sufficient time to follow the instructions given during the assessment.	0	5 (35.7%)	9 (64.3%)	0	1 (7.1%)	13 (92.2%)	-2.000	.046*
10-Telepractice can provide a satisfactory alternative to face-to-face assessment.	1 (7.1%)	4 (28.6%)	9 (64.3%)	0	2 (14.3%)	12 (85.7%)	-1.633	.102
11-Telepractice can allow easy access to assessment and therapy for children who stutter.	0	1 (7.1%)	13 (92.9%)	0	0	14 (100%)	-1.000	.317
12-I would prefer a traditional (face-to-face) assessment with the speech therapist	7 (50%)	2 (14.3%)	5 (35.7%)	13 (92.9%)	0	1 (7.1%)	-2.271	.023*

8.4 Summary of Findings

8.4.1 Feasibility Outcomes

30 sessions were carried out successfully using the telepractice model (whereby the TP system was used for carrying out assessment of 15 sessions and for observing children during face-to-face assessments in the other 15 sessions). There were instances where problems occurred, with four sessions being severely interrupted. However, none of these resulted in the termination of a session.

8.4.2 Validity and Reliability Outcomes

No significant differences were identified between the SSI-IV and SR scores obtained in the face-to-face and TP assessment environments, with the exception of %SS, which initially showed significantly different scores between face-to-face and telepractice; however, after conducting sensitivity analysis, the Wilcoxon signed rank test revealed no significant difference between %SS scores in TP and face-to-face sessions.

With regard to the level of agreement, initially, unacceptable levels of agreement were demonstrated for %SS by the Bland and Altman test. However, after conducting sensitivity analysis, a more satisfactory agreement level was achieved. Similarly, unacceptable levels of agreement of the SSI-IV subcomponents and SR were returned by the Bland and Altman analysis. However, when conducting percentage agreement analysis, results were satisfactory.

As for the reliability of face-to-face and TP ratings, all measures demonstrated satisfactory levels of intra- and inter-rater reliability, reflected by the ICC.

8.4.3 Acceptability Outcomes

Generally, parents and children exhibited high acceptability levels of satisfaction and held a positive view regarding telepractice pre- and post-assessment.

Possible explanations and implications of the results, in addition to whether the three proposed hypotheses were met or not, will be discussed in the following chapter.

9 Discussion

9.1 Introduction

The main aim of this study has been to investigate the feasibility, validity, reliability, and acceptability of conducting a stuttering assessment through the telepractice model of school-age CWS. In line with the aforementioned aims, three hypotheses were proposed.

The plan of this chapter is to discuss the results in accordance to these three research hypotheses. The results of each hypothesis will be summarised, interpreted, and discussed in line with the consistency of the study's results and the research literature. This is followed by a discussion of the benefits and challenges of applying the TP model and then the conclusion of this study. Considerations for SLTs will be outlined, with regard to adapting the TP model trialled in this study for the assessment of school-age CWS. The final section will consider the study's limitations and possible future research directions.

9.2 Discussion of Research Hypotheses

9.2.1 H1: Using the selected TP system, the TP stuttering assessment will be feasible and comparable to the face-to-face assessment.

All 30 sessions using the telepractice model (15 sessions leading and 15 sessions observing) were successfully carried out. Although there were instances where problems occurred, they did not prevent the sessions from continuing. Generally, the problems encountered were of two types: issues relating to the children and technological problems with audio and video quality. With regard to child-related issues, problems like the telepractice SLT losing sight of a child due to the child's change in position and a child speaking in an extremely quiet voice and being difficult for the TP and face-to-face SLTs to understand have been reported in previous telepractice studies (Fairweather et al., 2004; Hill et al., 2006; Theodoros et al., 2003; Waite et al., 2006; Waite et al., 2010).

As for technological issues, audio quality ratings were satisfactory for SLTs in most of the sessions and were rated as excellent in 80% of the sessions (see Section 8.3.1). However, visual quality was more of an issue, which is a commonly occurring problem in TP studies, especially those using low bandwidth internet (Carey et al., 2012; Constantinescu et al., 2011; Kully, 2002; Sutherland et al., 2017; Waite et al., 2010).

As mentioned earlier, none of the technological- or child-related problems encountered were sufficiently problematic to result in the termination of the sessions or an inability to communicate with the children and their parents. However, they did influence the scoring of the speech samples gathered via the TP system, thereby jeopardising the reliability and validity of stuttering assessment tasks scores. This will be examined in greater detail in the discussion of the following hypothesis.

In conclusion, the issues reported did not prevent the sessions from being carried out, making the chosen TP system feasible to carry out assessments. Most of the technical issues were related to visual quality, as the audio was stable in most of the sessions.

9.2.2 H2: Results from the assessment will present a high degree of clinically acceptable levels of agreement and reliability between the two environments, validating the TP modality for stuttering assessment.

To test this hypothesis, the validity and reliability of different stuttering measures, namely %SS, SSI-4 (total score and final severity ratings) and its subcomponents (%SS, duration, and physical concomitants), and SR were tested in TP and face-to-face sessions.

9.2.2.1 The reliability and validity of %SS via the TP model:

The first measure to be investigated was %SS, which revealed significant differences between the face-to-face and TP environments, as demonstrated by the Wilcoxon test. In addition, when further analysis of the agreement was conducted using the Bland and Altman method, the results were also unsatisfactory. This was clear, as the LOAs for the two conditions (face-to-face and TP) were between -4.33 and 5.97, which is considered to be a clinically unacceptable difference, as a difference of 4–6%SS may, for example, shift the severity level from mild to moderate. However, after inspecting the plot, we observed that the LOAs were wide due to nine speech samples belonging to seven participants, four of which had had sessions with a compromised internet connection. After conducting the sensitivity analysis (see Chapter 8), the LOA evidently improved when the assessments exhibiting technical difficulties were excluded, with %SS ranging from to -2.67% to 3.13%. It should be noted that, as suggested by previous research (Karimi et al., 2014a), the investigator's judgment is required to determine whether such inaccuracies may be acceptable in particular applications. This was a challenge in the present study, especially as there is no agreed-upon absolute numerical acceptable difference of %SS. Instead, we based our decisions on the idea that

%SS demonstrates acceptable levels of agreement, to the extent that it can be trusted to identify the size of the effect being sought. In other words, the investigator needs to generate statistical evidence that any effects found are due to changes in the independent variable/s and not the error variance of %SS. Therefore, based on clinical judgements and previous research (Bridgman et al., 2016; Carey et al., 2014; Karimi et al., 2013), a difference of approximately 3%SS was judged as insignificant, thus confirming that the assessment of %SS via TP is valid and reliable only when a reliable internet connection is available.

SLTs reported some difficulties in scoring %SS in some of the TP sessions, which may partially explain the discrepancy in the scores. One challenge was the absence of visual presentation in the sessions with compromised internet connection, which resulted in the judges relying only on the audio when scoring the recordings. Instances of pictures freezing and the loss of visual data occurred. However, to varying degrees, this technical difficulty has been commonly reported in TP assessment studies (Constantinescu et al., 2011; Sutherland et al., 2017; Waite et al., 2010). The severe loss of visual representation occurred in the case of four participants in this study, and it was these four participants that had the largest variations in LOA.

In a similar vein, in studies where the focus was to compare ratings of %SS in audio vs. audio-visual presentation, there was a mean increase of 18% in %SS scores when, compared to audio-only mode, samples of AWS were presented in audio-visual mode (O'Brian et al., 2015). The %SS was statistically significantly lower, by around 20%, when made from audio-only recordings of preschool CWS (Rousseau et al., 2008). In our study, on average, face-to-face ratings were slightly higher than TP ratings, and ratings of the samples of compromised video quality were all underestimated in the TP environment, compared to the face-to-face ratings, which may be the result of the loss of visual presentation. In addition, the judges of this study reported that, in some instances, they were unsure if it was a stuttering moment or a glitch, and, in some instances, where visual quality was compromised, they reported difficulty is measuring subtle stuttering moments (e.g. inaudible short blocks). However, this is not unique to the TP environment, as scoring the %SS is a challenge, even in traditional face-to-face studies (O'Brian et al., 2013). Moreover, in one instance, SLTs reported that it was difficult to score a sample due to the participant's low volume. However, this difficulty was reported by both the face-to-face and TP SLTs, in addition to also being reported in other traditional face-to-face studies (O'Brian et al., 2010).

In summary, it appears that, only when reliable internet is ensured, the TP model a valid and reliable method to assess %SS.

9.2.2.2 The reliability and validity of SSI-IV via the TP model:

The second measure of interest is the SSI-IV and the SSI-IV subcomponents (%SS, duration, and physical concomitants), along with the total score and the final severity rating.

The %SS subcomponent score agreement between face-to-face and TP environments was unsatisfactory, as the LOAs were between -2.18 and 2.32, and an approximate difference of two points is undesirable. However, it is thought that this poor agreement was not caused by the environment of the TP but rather the property of the test itself. Our results are comparable to our percentage agreement estimates of the SSI-IV reported in Chapter 5 and Davidow and Scott's (2017) study. A possible explanation for this unsatisfactory agreement, reflected in an approximately two sub-score difference, may be related to the sub-score range corresponding to the raw %SS, where the upper boundary of one sub-score range is close to the lower boundary of the next highest sub-score range (Davidow & Scott, 2017). In simpler terms, while two scorers may have small differences in raw %SS, they could end up with two different sub-scores. In an example from our data, a difference as small as 1%SS, where one scorer had a score of 2%SS and the other had a score of 3%SS, resulted in two different sub-scores (six and eight respectively) from the non-reader conversion table. On the other hand, in the case of higher scores of raw %SS, large differences between raw %SS may be masked when converted to scaled scores. In an example from our dataset, during a reading task, the face-to-face clinician scored 38.8%SS, whereas the TP clinician scored 24.3%SS, which is around a 14%SS difference. However, when converting to subscale scores, both ratings correspond to a score of nine. This raises the question of the usefulness of scaled scores and whether using raw data to track change is of greater benefit (Davidow & Scott, 2017).

The second subcomponent of the SSI-IV is the duration subcomponent, which is calculated by averaging the three longest stuttering moments. The Bland and Altman analysis revealed an unacceptable LOA. What is interesting here is that, when examining the plot for scaled score ratings, they either totally agreed or they differed by two points, which is one rank higher or lower on the conversion tables of the duration scores. However, as noted by Lewis (1996), it is thought that the raw counts do not reflect scaled scores. As seen in our data, a difference of as little as 0.11 seconds shifted the rating from ten to twelve points. On the other hand, a difference as large as 1.41 seconds only resulted in a two-point difference (from eight

to ten). Percentage agreements of exact scores within the two sub-score values were comparable to our reliability estimates of the SSI-IV and Davidow and Scott's (2017) study. This suggests that the poor agreement of the scaled duration sub-score may not be caused by the TP environment but rather by problems with the measurement itself.

The third subcomponent of the SSI-IV is the physical concomitants subcomponent, and, as in the %SS and duration subcomponents, LOAs were wide. In addition, percentage agreement indices were unsatisfactory.

This is not surprising or unique to this study, as these findings are in line with other studies that have investigated the agreement and reliability of the SSI. In Hall et al.'s (1987) study, the physical concomitants sub-score had the lowest inter-rater agreement, with a value of 54%. Moreover, in Davidow and Scott's (2017) study, SSI-IV revealed unsatisfactory inter-rater agreement, while, in Bakhtiar et al.'s (2010) study, where they examined the reliability of the Persian version of SSI-III, the least favourable result was in the physical concomitants component, with a value of 54% for inter-judge agreement. It can be shown that the discrepancies in the scoring may be linked to the quality of recording during the TP session, especially considering the subtle nature of some secondary features, and especially when, in this study, the majority of difficulties mentioned during the scoring were related to visual quality. However, it is thought that these differences cannot be fully attributed to the TP model. The usefulness of SSI's physical concomitants behaviour sub-score has always been questioned in the literature, as a result of its low reliability. This may be due to its subjective nature, as clear distinctions between each level of distraction are not provided (Riley, 2009; Tichenor, 2010). In other words, the way we differentiate a moment of stuttering with a distraction score of 1 ('not noticeable unless looking for it') or 2 ('barely noticeable to the casual observer') is not clear. The same is true of a score of 3 ('distracting') or 4 ('very distracting') since what is distracting to one observer may not be distracting to another (Tichenor, 2010). Further, it may be that some secondary features are more easily observed than others. Discrepancies between the results of the two assessment environments may also be due to methodological problems. Sometimes, camera angles were not satisfactory (see Section 9.5 below), and, if the child moved, this would result in a compromised view of the child's physical features accompanying the stuttering. This limitation has been documented in the TP literature in both paediatric and adult studies (Fairweather et al., 2004; Hill et al., 2006; Theodoros et al., 2003; Waite et al., 2006; Waite et al., 2010). Although the camera

angles were the same between the two recordings, the live exposure of the face-to-face SLT may have influenced the rating. In addition, this subcomponent encompasses the sum of ratings across different categories (distracting sounds, facial grimaces, etc.). As pointed out by Lewis (1996), totalling the judges' ratings across categories to calculate a single physical concomitants sub-score can obscure marked differences in judges' ratings on individual physical concomitant behaviours.

With regard to the total score, which is calculated by adding up the three components of the SSI, as for previous components, the LOA presented unsatisfactory results. This was not unexpected, as the total scores were derived from the sum of the components of the %SS, duration, and physical concomitants, which all posed unsatisfactory agreements using LOA, so it was likely that the accumulative differences would produce a different score. When comparing our findings of percentage agreement, unsatisfactory results were produced. However, this is not unique to this study, as such findings are in line with both Davidow and Scott (2017) and our reliability investigation outlined in Chapter 5.

In summary, results suggest that the TP model can be a valid and reliable method for assessing SSI-IV. However, due to the inherent problems of the SSI-IV itself, it is not recommended to use this model when the aim is to measure small changes within individual participants; in this case, SSI-IV should be interpreted with caution, which is reflected in the poor agreement that occurred when using the Bland and Altman method.

9.2.2.3 The reliability and validity of SR via the TP model

SR was the final measure investigated, and, as for the previous stuttering measures, SR failed to meet the clinical criteria for the Bland and Altman method comparison, as the LOAs were wide, ranging from -2.54 to 2.9 points difference, while the percentage level of the agreement level was 87.7%, within one scale value. This indicates that most ratings were within the acceptable difference limit of one scale value (Eve et al., 1995; O'Brian et al., 2004; Onslow, Andrews & Costa, 1990), which is in line with previous investigations into the reliability of SR. For example, a study by Karimi et al. (2014a) revealed an unsatisfactory LOA for SR. However, the percentage of agreement among the judges ranged from 86 to 95% of ratings being within the acceptable one score difference. In addition, O'Brian et al. (2004) also reported an excellent percentage of agreement between one experienced clinician and nine out of ten clients when using a 9-point SR scale: 78% of scores were within one scale value of

each other. Our percentage agreement calculations may have been higher than O'Brian et al.'s (2004) due to the limited number of judges, which may have led to higher ratings.

Unlike the %SS, in exploring the Bland and Altman plot, SR seemed reliable, even in the case of compromised internet quality, which may be due to the different constructs of %SS and SR, as, in contrast to SR, %SS involves the discrete counting of syllables. In this regard, SR requires an overall impression that could be maintained regardless of the loss of data that may occur as a result of a technological glitch. It has been documented that, unlike %SS, SR scales are not affected by the presentation mode (i.e. audio-only/audio-visual); for example, in Rousseau et al. (2008), the judges' scores on a 9-point scale did not differ across the two presentation modes. This could indicate that SR may be a more suitable measure than %SS to use during a TP assessment, especially since technological glitches are unavoidable. However, complete dependency on SR alone for the assessment of overt features of stuttering may be insufficient (O'Brian et al., 2004).

It can be concluded that the lack of agreement of SR seen in LOA was not due to any deficits in the information provided by the TP environment. Instead, it was related to the SR measure itself, as the results of this study were similar to traditional face-to-face research incorporating SR.

In summary, our findings support the proposed hypothesis. However, this is only true when reliable internet is ensured, especially in the case of %SS. The TP model seemed to be a valid and reliable option for a stuttering assessment, as, when the differences in assessment results were found, they were similar to those reported in the traditional face-to-face literature. It should be noted that one must be aware of the assessments' shortcomings as highlighted in Chapter Two. Nevertheless, these measurements remain central to clinical practice (although other measures targeting other aspects of stuttering should be considered), as they help to shape decisions at every step of the clinical process, including assessment, intervention, termination of treatment, and follow-up (McCauley,1989).

9.2.3 H3: Children and their parents will feel positive about the potential of using telepractice for stuttering assessment, satisfied with the TP assessment of their stuttering, and willing to use the TP system for stuttering therapy in the future.

To our knowledge, there has been no investigation into client or parent satisfaction of stuttering assessment services delivered via TP. Instead of only collecting views post-assessment, as most TP studies have done, views were gathered both pre- and post-assessment, as it was thought that considering preconceptions about stuttering assessment services may have an impact on the willingness of clients to take up this service. It was therefore decided that this should be an essential part of the evaluation of patient perception and satisfaction in the proposed TP assessment (Sharma et al., 2013).

The post-assessment responses of parents and children indicated that they both had a generally positive experience of TP assessment. Looking at the pre-assessment responses, the majority expressed a willingness to use this model to receive services, which is encouraging, although noting that they had not tried it before. Moreover, more positive opinions regarding the service were noted post-assessment, as both parents and children reported that the TP model did not differ from the face-to-face model, and, most interestingly, they did not prefer the traditional face-to-face assessment over the TP model. Such a positive trend is promising, and the fact that only one exposure to such a service elicited positive responses is indicative of the parents' and children's willingness and openness to utilise the TP model, even though they have had limited experience with it.

As mentioned earlier, this study revealed high levels of satisfaction in both parents and children when using the TP assessment model, which is in line with research that has incorporated this model for assessing different speech and language disorders in children. In a study by Ciccia et al. (2011), the feasibility of conducting SLT screenings via Skype was investigated, finding high levels of satisfaction from the parents. In fact, parents indicated a preference for accessing these services via Skype rather than having to travel to complete assessments. Recently, a study by Sutherland et al. (2107) examined the feasibility of conducting language assessments of children via TP and highlighted the positive attitudes of both the parents and their children, who stated that they felt comfortable using the TP system. In the field of TP and stuttering, similar trends were observed in TP stuttering treatment studies (Bridgman et al., 2016; Carey et al., 2012).

In the present study, children reported feeling relaxed during the TP session, particularly when compared to a typical face-to-face session. Although this seems intriguing, such a finding should be considered with caution, as the goal of a stuttering assessment is to reflect everyday life, and the child being more relaxed in this environment may lead to the omission of crucial information. On the other hand, one child reported feeling more anxious when compared to the face-to-face session with the SLT, so this assessment may have been overestimated. However, a similar problem can be encountered in face-to-face assessments, especially those carried out post-treatment, as such assessments may be biased, either because of discriminated learning to the clinic (Wilkie & Beilby, 1996) or the improvement of fluency potentially being due to the fact that the client is already familiar with the clinician and hence more relaxed when compared to the pre-assessment, and the improvement is therefore more of a placebo effect (Boberg & Kully, 1985). Either way, it seems that the stuttering variability of clients is unavoidable, as it is a central feature of stuttering (Constantino et al., 2016).

Human factors should also be kept in mind while evaluating a TP service (Waite et al., 2010). One dimension of this is building rapport, which is a concern in the literature, as it has been reported that rapport building may be a challenge (Hines et al., 2015; May & Erickson, 2014; Tucker, 2012), especially for young children (Fairweather et al., 2004). However, our data suggests otherwise, as comments from the two therapists and the children indicated that this was not an issue. Children reported feeling that communicating in this medium was comfortable and easy, and they felt less pressure, given the physical absence of the therapist. Moreover, the ease of communication may also be attributed to the children's enthusiastic response to technology, which aided their focus and attention (Fairweather et al., 2004). One vital factor that supports the success of stuttering management is the interpersonal relationship between the clinician and the client (Cooper, 1966), Murphy and Fitzsimons (1960), for example, stated that 'much of the past therapeutic success of speech clinicians with stutterers can be attributed to the undefined emotional relationship that existed between clinician and stutterer, rather than to the specific speech correction technique employed' (p.8). In this regard, being able to build rapport with the client during assessment (which is the first encounter of therapy) may support a positive interpersonal relationship between the clinician and the client.

On the other hand, parents might hold a different point of view when evaluating the service. In one instance, in the post-assessment questionnaire, a mother reported being concerned

about building rapport with a distant clinician, and she thought that it would be better to have face-to-face session initially in order to overcome that barrier. However, when her child was asked to describe his experience, he reported feeling very comfortable and, most importantly, he commented that he was able to connect and communicate better with the clinician as a result. However, this comment was only made by one parent in this study, so, in the future, it would be interesting to compare the views of the children and their parents in depth, especially given that the consent to take up such a service must be gained from the parent of a young child.

In summary, our findings support the third hypothesis, as parents and children exhibited high levels of satisfaction and held positive views of TP, both pre- and post-assessment.

9.3 Benefits, Drawbacks, and Clinical Implications of the TP Model in the Assessment of School-Age CWS

Based on our findings, it appears that TP is a promising approach for providing stuttering assessment, resulting in similar outcomes to the traditional face-to-face approach. In this investigation, we were successful in conducting TP sessions with limited infrastructure. More specifically, we were able to conduct stuttering assessments using a laptop, webcam, and free video conferencing software, simply by connecting to the internet. The study found that there was no statistical differences between TP and face-to-face administration on the outcome measures tested, namely %SS, SR, and SSI-IV, which provides support for the use of the TP model to assess school-age CWS using these measures. Furthermore, although not addressed statistically, the model was successful in discussing the attitudes and impacts of stuttering on the participants' daily lives. These findings have implications at both clinical and research levels.

From a clinical point of view, we were able to collect speech samples of children interacting with their parents while the TP SLT was observing silently. Thus, the TP model may facilitate collecting beyond-clinic speech samples. This is a part of the stuttering assessment in clinical contexts that enables us to gain a perspective of a child's stuttering in different situations, which may also be helpful in assessing children who are shy or reluctant to interact with the clinician, especially in their first sessions (Vogel et al., 2015). Furthermore, such beyond-clinic samples are an integral part of evaluating the progress of therapy programmes for CWS, such as the PCIT therapy program (Kelman & Nicholas, 2017). This may also be a helpful avenue in maintaining regular follow-ups and assessing the maintenance of therapy for

children who cannot go to the clinic. In other words, given the obstacles in Saudi Arabia, i.e. the limited number of fluency specialists, sparse areas, and unequal distribution of therapists, this model could increase access to specialists in the field of stuttering assessments, consultations, and treatments, especially for this age group, given that technology is part of their daily lives (Lowe, O'Brian & Onslow, 2013). The TP model may allow children to access services from the convenience of their own homes, eliminating the need to travel, which could increase their commitment to attending sessions and decrease missed appointments, thereby creating better therapy outcomes (Lowe, O'Brian & Onslow, 2013; O'Brian et al., 2003). In addition, the success of some types of therapy is dependent on how much the parents are involved in the therapeutic process (in PCI, both parents must attend the therapy sessions), and the flexibility of this model may help with the consistent involvement of family members. On the other hand, for older children, specifically adolescents, the TP model could encourage independence and self-management, both of which are attributes linked to improved outcomes (Bothe et al., 2006), as children can have a session with the SLT alone without their parents being present (Lowe, O'Brian & Onslow, 2013).

In addition to the aforementioned benefits of the TP model in clinical contexts, telepractice can also be utilised in the assessment of stuttering in research contexts. The outcome measures evaluated in this study (%SS, SR, and SSI-IV) are commonly used in stuttering research, and researchers may be able to administer these outcome measures via TP in research studies. Given its flexibility, and eliminating the need to travel, TP may help in recruiting and increasing the compliance of participants in clinical trials, thus increasing sample sizes. Having an adequate sample size in a research study is vital for meeting the desired statistical power. This can be a challenge in stuttering research due to the relatively low prevalence of the disorder (Jones et al., 2002). Further, the TP model can help the researcher ensure that the quality of the recorded speech sample is adequate during the time of recording, in addition to also helping the researcher monitor the quality of parental engagement of the child, specifically ensuring that the parent is involved with the child in a truly conversational dialogue. Finally, the recording will be immediately available for backing up and re-measuring, if required, which further adds to its practicality (Vogel et al., 2015).

In this study, in TP-led sessions, the face-to-face SLT acted not only as observer but also as a facilitator. However, tasks and activities during the session did not require assistance from the face-to-face SLT, as the face-to-face SLT's role did not go beyond helping to establish the

video conference call and helping to reconnect with the TP clinician in case the connection was compromised. This may indicate that a stuttering assessment via TP can be conducted without the presence of a specialist. In this regard, the parent or, in the case of older children, the child him/herself could manage the session according to the instructions of the TP clinician.

In summary, our study suggests that, when compared to traditional face-to-face assessment, the TP stuttering assessment of school-age children is feasible, valid, and a reliable method of assessment. It has also been shown that, for the TP assessment, the use of simple equipment was appropriate for carrying out the assessment and to rate samples. Further, its benefits may extend to both clinical and research contexts. The TP system chosen for this study was successful in allowing the assessments to be carried out in terms of consultation, where the clinician and the participants were able to interact and communicate adequately via the chosen system. It also provided acceptable quality, in terms of audio and video recordings, for the assessment of %SS, SSI-4, and SR. Moreover, overall, the majority of children and their parents in the TP environment felt satisfied with the TP delivery model, which further supports the implementation of this method.

However, regardless of the potential benefits and findings of this study, it is also important to discuss the challenges encountered. The most prominent limitation in using the TP model for the assessment of CWS related to technological problems, with glitches occurring to varying degrees during the sessions. All sessions sufficiently communicated with and gathered information from clients, and the difficulties only arose when rating samples for stuttering severity measures. The following section will discuss these technological challenges and what impact they had on the assessment.

Generally, the quality of both audio and video during video conferencing is dependent on the bandwidth. A low bandwidth connection usually compromises the visual quality rather than the audio quality, which happened in our case, where image quality was the most compromised feature during the video conferencing. In this regard, SLTs encountered difficulties, such as the pixilation, freezing, and skipping of images, as well as a total loss of visuals, which made scoring more difficult when compared to the samples taken in face-to-face situations. However, these difficulties are not unique to this study, as many TP studies, especially those depending on low bandwidth video conferencing, also reported such

difficulties (Carey et al., 2012; Constantinescu et al., 2011; Kully, 2002; Sutherland et al., 2017; Waite et al., 2010).

The audio quality was far less affected and was generally satisfactory. However, there were instances of audio skipping and delays in audio signal transmission. During the session, delays in the audio transmission were easily adjusted, and the PhD candidate followed recommendations from previous research (Kully, 2002), as the participant was instructed to wait for two seconds before responding to make up for the delay. However, in some instances, during the scoring of the speech samples gathered, SLTs reported that, with participants exhibiting subtle silent blocks without struggling behaviours, it was confusing to differentiate whether this was a stuttering moment or whether the patient was simply adapting to the audio delay.

Regardless of these shortcomings, SLTs reported that the compromised image quality did not affect the flow of the assessment nor their ability to build rapport. Parents and children also reported high ratings for the audio and video quality, meaning that the majority of people were satisfied. However, as the quality influenced the scoring, when compared to face-to-face samples, the SLTs reported that they were less confident in their ratings. Moreover, additional difficulties in viewing the children during the video conferencing occurred because of participant-related factors, such as the child moving or speaking in an extremely low voice due to shyness or as a form of avoidance. However, these factors affect both face-to-face and TP situations (Waite et al., 2010).

During the scoring of the TP sessions, if there was confusion between a technological anomaly and a stuttering moment, this was not counted in the percentage of stuttered syllables, which may partly explain the underestimation of the TP ratings in %SS, especially in the sessions where severe technical issues were encountered. As Valentine (2014) suggested, in the case of frequent interruptions to the connection during the session, the SLT and participant would log-off and re-establish the connection, which would clearly improve the internet connection.

As mentioned earlier, although the audio and video quality was sufficient to carry out the assessment during video conferencing, SLTs reported that scoring the recorded samples via the TP system was challenging at times, especially when counting subtle stuttering moments, e.g. silent blocks or mild facial grimaces, which is thought to have been due to the reliance on low bandwidth for video conferencing. However, in the future, the availability of reliable,

high bandwidth internet will offer better audio and video clarity, thus enhancing the experience of video conferencing (Lowe, O'Brian & Onslow, 2013). In case technological glitches occur during the session, SLTs can record the session using alternative methods, such as downloading recording software at the participant's end. After the session is completed, the participant can share the recording with the SLT via e-mail. This way, the session will be carried out via TP, and at the same time, a sample that does not rely on the internet is available for the SLT.

Other than technological challenges, SLTs observed that there were differences in establishing eye contact, specifically describing the fact that eye contact felt somewhat unnatural in the TP environment when compared to face-to-face sessions. Sometimes, participants looked up at the camera on top of the screen rather than looking directly at the screen to establish eye contact, which resulted in a gaze that appeared to be looking up. This was sometimes confusing to the SLTs, as it could be perceived as avoidance behaviour. Therefore, the SLT should try to distinguish whether the lack of eye contact is a technical issue or a clinical one, i.e. a feature observed in stuttering (e.g. avoidance) (see Section 9.5 for some suggestions for overcoming the aforementioned challenges).

In summary, the current investigation has illustrated that school-age CWS can access assessment services using this model, and technical problems for both the clinician and the participant can be dealt with. However, as this study was exploratory in nature, generalisations are not possible and further investigation is necessity. In this sense, the present study has allowed us to conclude that further research in this area is worthwhile.

9.4 Conclusion

This study has suggested that conducting a stuttering assessment via TP for school-age CWS is feasible, reliable, and valid. Although the TP model posed some challenges for the execution of the assessment, mainly technological challenges, high levels of satisfaction were recorded by the parents and children alike. The TP system and equipment chosen in this study have provided a basis for the delivery of TP stuttering assessment in a clinical setting, whereby addressing the access barriers present for school-age CWS in Saudi Arabia. These results are preliminary, but they can be seen as a building block for future research in TP assessment studies for CWS.

9.5 Considerations for SLTs Applying the TP Model for the Assessment of Stuttering

In light of the observations during our study and recommendations from previous studies, this section outlines some considerations that should be taken into account before, during, and after conducting a stuttering assessment through TP in order to help enhance the TP assessment experience.

Before deciding to conduct a stuttering assessment via TP, the SLT and the parents of the child should engage in a discussion to determine whether TP is a suitable option. This discussion involves the SLT ensuring that the materials required (i.e. laptop, webcam, audio speakers, microphone, and internet connection) are available to the parents seeking the service. Once this is done, as suggested by Galvan, Case and Todd (2014), it is helpful to conduct a brief session to test the internet connection and to ensure that the parents and children are comfortable with the dynamic. In addition, this brief session also helps the SLT to ensure that the space parents have chosen to conduct the session is suitable for the assessment, is quiet, and has no distractions, in addition to helping them to train the child on the optimal positioning for the speech sample recording. Once all this is determined, the child is scheduled for a TP stuttering assessment session. In our study, we mentioned that eye contact was sometimes felt to be unnatural in the TP environment, where participants looked up at the camera on the top of the screen rather than looking directly to establish eye contact. To avoid this problem, the SLT can instruct the child to look at the screen so as to ensure eye contact during this introductory session.

During the assessment, it is recommended that the SLT investigates whether the TP environment has an effect on the child. In our study, children reported feeling more relaxed than usual due to the lack of the physical presence of the SLT in the same room. However, compared to a typical face-to-face session, high levels of anxiety were also reported. Understanding these feelings will help the SLT to create a more reflective assessment of the child in his/her day-to-day life.

Even if a test of the internet connection has been carried out prior the assessment, technological glitches are unavoidable, even if only to a minimal degree. In our study, we found that these glitches did not interfere with building rapport, carrying out interviews, or having conversations with the parents or their children. We tried to adjust to these difficulties by following procedures used in earlier research. In this regard, as suggested by Kully (2002), we adapted to delays in audio transmission by waiting two seconds before responding. When

the audio cut out, we made sure to request clarifications from participants and their parents, and we instructed them to do the same. Further, following Valentine's (2015) recommendation, SLTs and participants would log-off and re-establish the connection if there were frequent interruptions to the connection during the session, which worked in improving the internet connection.

The real downside we observed was that such technological glitches interfered with the quality of the speech sample recorded via the video conferencing software. Therefore, we suggest recording the session using a free video recorder software downloaded by the participant (e.g. https://www.nchsoftware.com/capture/index.html), which is not influenced by the quality of the internet connection. In this sense, the parent or child (in the case of older children) is requested to turn on the recorder during the session and then share the recording with the SLT by sending it via email after the session so that he/she can measure stuttering severity.

On a positive note, we found that some features of the chosen video conference software were helpful. For example, the whiteboard feature (see Chapter 7) permitted the presentation of pictures and reading material during sample speech recordings. In addition, the whiteboard feature allowed some activities to be carried out to maintain the child's interest and to build rapport, as the child was able to draw on the screen. The SLTs also asked some participants to draw how they represented their stuttering, which is a method used by clinicians to understand the children's views of their stuttering. Further, in some instances, SLTs presented the iceberg model to children by drawing on the whiteboard what appears on the participant's screen.

Another feature we found helpful is the picture-in-picture feature, which is available for both the SLT and the participant, whereby a small screen appears on the lower corner of the monitor so that the participant can see him/herself. This feature enhanced rapport building, as participants found it fun to view themselves on screen. However, in one instance, in which the participant was a younger child, this feature was distracting, so the SLT refrained from using it. Such a feature may go beyond this assessment, as it could be helpful during training on speech restructuring techniques, as the child could monitor him/herself during training.

After the session, it is suggested that the SLT should seek the opinions of both the parents and their children, as we found that, in some instances, a child can be excited about the TP method, whereas parents believe that it is better to have a face-to-face session and to try TP at

a later stage. Thus, understanding both views will help to balance the service preferences of both the parents and their children (Crais, Roy & Free, 2006; Eriks-Brophy et al., 2008; Taylor et al., 2014).

9.6 Limitations and Future Directions

There are several limitations associated with this study that should be taken into account in future research in order to provide a broader view of the feasibility, validity and reliability, and acceptance of TP stuttering assessment of school-age children who stutter. These limitations, along with future suggestions, are addressed next:

- Although the outcome of this study is of value, it is preliminary in nature, encompassing a limited number of SLTs (two) assessing and scoring samples, in addition to being mainly conducted by the principal investigator, who is a PhD candidate. Therefore, the study may be prone to researcher bias. A replication of this study using larger pools of SLTs and a larger cohort of patients will provide further evidence of the efficacy and efficiency of this model.

- This investigation was conducted in two rooms within the same hospital building, so it is a simulation of what it may be like in home environments, health units, or schools. Therefore, testing this model in each of these settings is the next logical step, which will provide further evidence of the feasibility of stuttering assessments in real-world settings, highlighting the realistic benefits and drawbacks of conducting stuttering assessments in each setting.
- In this investigation, we used one video conferencing platform, WebEx. It may be worth carrying out assessments using other commonly used software programs, such as Skype, as they may result in different outcomes in terms of feasibility, audio video quality, and acceptance among children and their parents.
- In this study, we used a software-based TP system (WebEx) with consumer-grade equipment (a laptop, webcam, and broadband internet) to conduct and record stuttering assessment sessions. This offers flexibility, which leads to increased access to specialist assessments at a minimal coast, with tools that are widely available to the community. However, when using such platforms, we encountered technological challenges, especially when attempting to maintain image quality. Although this did not prevent assessments from being carried out, it caused problems when rating

samples. In this regard, incorporating asynchronous methods may help to overcome poor audio and image quality in instances of compromised internet speed (Waite et al., 2006; Waite et al., 2010; Waite et al., 2012). Specifically, the TP SLT could carry out the assessment online to gather information, and then the session could be recorded using high-quality video recording independent from the internet (Waite et al., 2006; Waite et al., 2010), which may improve the accuracy of scoring.

- The outcome measures of this study (%SS, SR, an SSI-IV) were primarily overt measures. It would be interesting to determine whether QoL measures and attitudinal scales, such as the CAT and OASES, can be reliably completed via TP, as they may pose different challenges to those encountered in our study. It may be speculated that it would be less problematic to apply such measures via TP as they are self-administered and non-reliant on SLT judgement; however, this needs formal investigation. We have limited our study to overt measures because we were interested in whether SLTs can see, hear, and complete sessions over TP with assessment measures, just as the sessions would be completed in a face-to-face setting. However, investigating the reliability and validity of applying QoL and attitudinal scales via TP addresses a different question and thus requires a different methodology and study design (Nelson & Palsbo, 2006); in this sense, the question such a perspective poses is whether the same client–clinician relationship is established and whether rapport can be built in order for patients to complete the diagnostic interview as a whole. It is thought that this was partially addressed when seeking opinions of children and their parents regarding the TP assessment.
- Our sample mainly consisted of children living in the city of Riyadh (27 out of 30); thus, it would be interesting to implement this service delivery model in rural populations, as this may produce different results with regard to acceptability and readiness. In addition, infrastructure, particularly internet connectivity, could be more of an issue in these areas.
- The sample in this study consisted of CWS who are in their school years, so it would be interesting to implement this model with preschool children who stutter, especially as younger children may pose a greater challenge than those in our sample. If proven successful, this could provide an avenue for the early identification of stuttering by increasing access to preschool stuttering assessment services, which would result in a higher potential for recovery, leading to the amelioration of the consequences of

stuttering that are exhibited throughout one's lifespan, as discussed in Chapter One.

- We have investigated the acceptability and satisfaction of the TP model by using quantitative methods, but it may be more beneficial to incorporate qualitative research in order to explore the views of the participants and their parents in-depth and to fully understand the barriers and facilitators of TP. In addition, most of the parents who participated were mothers (27 out of 30), so it would be interesting to involve both mothers and fathers and to compare their views, as they might hold different opinions. Moreover, although the SLTs were asked to rate sessions in term of audio and image quality, investigating SLTs' views on more aspects of TP would lead to a better understanding of the application of the TP method among SLTs.
- Given the conservative nature of Saudi Arabian culture, it was anticipated that we may have encountered difficulties interacting with participants during the TP method, especially as they were aware that they were being video recorded, which can be a sensitive matter in such cultures. However, this was not the case at all, as building rapport and carrying out dynamic discussions was not a problem, perhaps because these people consented to be part of the study in the first place, so they had no issues with being recorded. However, the real challenge in our study was recruiting participants, especially older school-age females, as the study required the recording of the video conference. It should be noted that ten families that were approached declined the offer to participate in the study, which was mainly due to the fact that they felt their children would become exposed and may be recognised by seeing an SLT via TP and being recorded, as opposed seeing an SLT behind closed doors. This raises the question of whether this model is suitable for such a culture. Therefore, it is thought that investigating the acceptability of TP among Saudi stutterers in general would provide a better insight into their willingness to take up such a service. However, it seems like Saudi Arabia is heading in a new direction, given that the Saudi Crown Prince has announced the 'Saudi Vision 2030' plan to diversify its economy, develop the public service sectors, including the health sector, and transform Saudi Arabia into a more open and modern 21st-century society. In accordance with this programme, the MOH has expanded its telemedicine services by launching the 'Seha' application (see Section 6.2), which provides visual medical consultations. As mentioned in Chapter 6, such investments can certainly be

considered an endorsement of the validity, effectiveness, and future of telepractice within the Saudi healthcare system.

References

Abou Ella, M., Saleh, M., Habil, I., El Sawy, M. & El Assal, L. (2015) Prevalence of stuttering in primary school children in Cairo-Egypt. *International Journal of Speech-Language Pathology*. 17 (4), 367–372.

Adams, M.R. (1990) The demands and capacities model I: Theoretical elaborations. *Journal of Fluency Disorders*. 15 (3), 135–141.

Alameer, M., Meteyard, L. & Ward, D. (2017) Stuttering generalization self-measure: Preliminary development of a self-measuring tool. *Journal of Fluency Disorders*. 53, 41-51

Al-Jadid, M.S. (2013) Disability in Saudi Arabia. *Saudi Medical Journal*. 34 (5), 453–460.

Alkabba, A.F., Hussein, G.M., Albar, A.A., Bahnassy, A.A. & Qadi, M. (2012) The major medical ethical challenges facing the public and healthcare providers in Saudi Arabia. *Journal of Family and Community Medicine*. 19 (1), 1.

Al-Khaledi, M., Lincoln, M., McCabe, P. & Alshatti, T. (2017) The Lidcombe Program: A series of case studies with Kuwaiti preschool children who stutter. *Speech, Language and Hearing*. 1–12.

Alpermann, A., Huber, W., Natke, U. & Willmes, K. (2010) Measurement of trained speech patterns in stuttering: Interjudge and intrajudge agreement of experts by means of modified time-interval analysis. *Journal of Fluency Disorders.* 35 (3), 299–313.

Alpermann, A., Huber, W., Natke, U. & Willmes, K. (2012) Construct validity of modified time-interval analysis in measuring stuttering and trained speaking patterns. *Journal of Fluency Disorders*. 37 (1), 42–53.

Alquraini, T. (2010) Special education in Saudi Arabia: Challenges, perspectives, future possibilities. *International Journal of Special Education*. 25 (3), 139–147.

Alshahrani, H.A. (2016) A brief history of the internet in Saudi Arabia. *TechTrends*. 60 (1), 19–20.

Alyemeni, M. (2014) Five-year program to transform healthcare delivery in Saudi Arabia. *Saudi Arabia Ministry of Health.* Available from: http://www.salud-e.cl/wp-content/uploads/2013/08/ saudi+arabia+e-health+strategy.pdf [Accessed 18th June 2017].

American Psychiatric Association. (2013) *Diagnostic and Statistical Manual of Mental Disorders*. Fifth edition. American Psychiatric Publishing, Arlington, VA.

American Speech-Language-Hearing Association. (1985) *1985 National Colloquium on Underserved Populations Report*. 27 (11), 31–35.

American Speech-Language-Hearing Association. (2008) *Schools Survey Report: Workforce and Work Conditions Trends 2000–2008*. Rockville, MD: American Speech-Language-Hearing Association.

American Speech-Language-Hearing Association. (2013) *Telepractice*. Available from: https://www.asha.org/Practice-Portal/Professional-Issues/Telepractice/ [Accessed 10th May 2015].

Andrews, G., Morris-Yates, A., Howie, P. & Martin, N. (1991) Genetic factors in stuttering confirmed. *Archives of General Psychiatry*.

Armson, J. & Kiefte, M. (2008) The effect of SpeechEasy on stuttering frequency, speech rate, and speech naturalness. *Journal of Fluency Disorders*, 33 (2), 120–134.

Arnott, S., Onslow, M., O'Brian, S., Packman, A., Jones, M. & Block, S. (2014) Group Lidcombe Program treatment for early stuttering: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research*. 57 (5), 1606–1618.

Atkinson, G. & Nevill, A.M. (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*. 26 (4), 217–238.

Ayre, A. & Wright, L. (2009) WASSP: An international review of its clinical application. *International Journal of Speech-Language Pathology*. 11 (1), 83–90.

Bainbridge, L.A., Stavros, C., Ebrahimian, M., Wang, Y. & Ingham, R.J. (2015) The efficacy of stuttering measurement training: Evaluating two training programs. *Journal of Speech, Language, and Hearing Research*. 58 (2), 278–286.

Bakhtiar, M. & Packman, A. (2009) Intervention with the Lidcombe Program for a bilingual school-age child who stutters in Iran. *Folia Phoniatrica et Logopaedica*. 61 (5), 300–304.

Bakhtiar, M., Seifpanahi, S., Ansari, H., Ghanadzade, M. & Packman, A. (2010) Investigation of the reliability of the SSI-3 for preschool Persian-speaking children who stutter. *Journal of Fluency Disorders*. 35 (2), 87–91.

Bartlett, J.W. & Frost, C. (2008) Reliability, repeatability and reproducibility: Analysis of measurement errors in continuous variables. *Ultrasound in Obstetrics & Gynecology*. 31 (4), 466–475.

Barnhart, H.X., Haber, M.J. & Lin, L.I. (2007) An overview on assessing agreement with continuous measurements. *Journal of Biopharmaceutical Statistics*. 17 (4), 529–569.

Batterham, A.M. and George, K.P. (2003) Reliability in evidence-based clinical practice: A primer for allied health professionals. *Physical Therapy in Sport*. 4 (3), 122–128.

Beal, D.S., Gracco, V.L., Brettschneider, J., Kroll, R.M. & De Nil, Luc, F. (2013) A voxel-based morphometry (VBM) analysis of regional grey and white matter volume abnormalities within the speech production network of children who stutter. *Cortex*. 49 (8), 2151–2161.

Beal, D.S., Gracco, V.L., Lafaille, S.J. & De Nil, Luc, F. (2007) Voxel-based morphometry of auditory and speech-related cortex in stutterers. *Neuroreport*. 18 (12), 1257–1260.

Bland, J.M. & Altman, D.G. (1999) Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*. 8 (2), 135–160.

Bland, J.M. & Altman, D.G. (2003) Applying the right statistics: Analyses of measurement studies. *Ultrasound in Obstetrics & Gynecology*. 22 (1), 85–93.

Bland, J.M. & Altman, D.G. (1990) A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Computers in Biology and Medicine*. 20 (5), 337–340.

Bland, J.M. & Altman, D. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 327 (8476), 307–310.

Blood, G.W. & Blood, I.M. (2004) Bullying in adolescents who stutter: Communicative competence and self-esteem. *Contemporary Issues in Communication Science and Disorders*. 31 (31), 69–79.

Bloodstein, O. (1995) *A Handbook on Stuttering.* San Diego, CA: Singular Publishing Group.

Bloodstein, O. & Bernstein Ratner, N. (2008) *A Handbook on Stuttering.* 6th ed. Clifton Park, NY: Delmar.

Blumgart, E., Tran, Y. & Craig, A. (2010) An investigation into the personal financial costs associated with stuttering. *Journal of Fluency Disorders*. 35 (3), 203–215.

Boberg, E. & Kully, D. (1985) *Comprehensive Stuttering Program: Clinical Manual/Cassette*. College Hill Press.

Boberg, E. & Kully, D. (1994) Long-term results of an intensive treatment program for adults and adolescents who stutter. *Journal of Speech, Language, and Hearing Research*. 37 (5), 1050–1059.

Boey, R.A., Van de Heyning, P.H., Wuyts, F.L., Heylen, L., Stoop, R. & De Bodt, M.S. (2009) Awareness and reactions of young stuttering children aged 2–7 years old towards their speech disfluency. *Journal of Communication Disorders*. 42 (5), 334–346.

Bothe, A.K., Davidow, J.H., Bramlett, R.E. & Ingham, R.J. (2006) Stuttering treatment research 1970–2005: A systematic review incorporating trial quality assessment of behavioral, cognitive, and related approaches. *American Journal of Speech-Language Pathology*. 15 (4), 321–341.

Boyle, C.A., Boulet, S., Schieve, L.A., Cohen, R.A., Blumberg, S.J., Yeargin-Allsopp, M., Visser, S. & Kogan, M.D. (2011) Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics*. 127 (6), 1034–1042.

Brady, A. (2007) Moving toward the future: Providing speech-language pathology services via telehealth. *Home Healthcare Now*. 25 (4), 240–244.

Braun, A.R., Varga, M., Stager, S., Schulz, G., Selbie, S., Maisog, J.M., Carson, R.E. & Ludlow, C.L. (1997) Altered patterns of cerebral activity during speech and language production in developmental stuttering. An H2 (15) O positron emission tomography study. *Brain: A Journal of Neurology*. 120 (5), 761–784.

Brennan, D.M., Georgeadis, A.C., Baron, C.R. & Barker, L.M. (2004) The effect of videoconference-based telerehabilitation on story retelling performance by brain-injured subjects and its implications for remote speech-language therapy. *Telemedicine Journal & e-Health*. 10 (2), 147–154.

Bridgman, K., Onslow, M., O'Brian, S. & Block, S. (2011) Changes to stuttering measurement during the Lidcombe Program treatment process. *Asia Pacific Journal of Speech, Language and Hearing*. 14 (3), 147–152.

Bridgman, K., Onslow, M., O'Brian, S., Jones, M. & Block, S. (2016) Lidcombe Program webcam treatment for early stuttering: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research*. 59 (5), 932–939.

Brundage, S.B., Bothe, A.K., Lengeling, A.N. & Evans, J.J. (2006) Comparing judgments of stuttering made by students, clinicians, and highly experienced judges. *Journal of Fluency Disorders*. 31 (4), 271–283.

Brutten, G.J. (1985) Communication attitude test: Form A. *American Journal of Speech-Language Pathology*. 6 (4), 73.

Brutten, G.J. & Dunham, S.L. (1989) The Communication Attitude Test: A normative study of grade school children. *Journal of Fluency Disorders*. 14 (5), 371–377.

Brutten, G.J. & Vanryckeghem, M. (2007) *Behavior Assessment Battery for School-Age Children Who Stutter*. Plural Publishing.

Buchanan, T. (2002) Online assessment: Desirable or dangerous? *Professional Psychology: Research and practice*. 33 (2), 148.

Burgess, L.P., Holtel, M.R., Syms, M.J., Birkmire- Peters, D.P., Peters, L.J. & Mashima, P.A. (1999) Overview of telemedicine applications for otolaryngology. *The Laryngoscope*. 109 (9), 1433–1437.

Bryson, S.E. (1996) Brief report: Epidemiology of autism. *Journal of Autism and Developmental Disorders*. 26 (2), 165–167.

Canadian Association of Speech Language Pathologists and Audiologists. (2006) *Position paper on the use of telepractice for CASAPLA Speech-Language Pathologists and Audiologists*. Available from: https://www.sac-oac.ca/professional-resources/resource-library/sac-position-paper-use-telepractice-sac-s-lps-and [Accessed 3rd June 2017].

Cardell, E.A. (2010) Review of: Communication attitude test for preschool and kindergarten children who stutter (KiddyCAT). *International Journal of Disability Development and Education*. 57 (2), 236–238.

Carey, B., O'Brian, S., Lowe, R. & Onslow, M. (2014) Webcam delivery of the Camperdown Program for adolescents who stutter: A Phase II trial. *Language, Speech, and Hearing Services in Schools*. 45 (4), 314–324.

Carey, B., O'Brian, S., Onslow, M., Block, S., Jones, M. & Packman, A. (2010) Randomized controlled non-inferiority trial of a telehealth treatment for chronic stuttering: The Camperdown Program. *International Journal of Language & Communication Disorders*. 45 (1), 108–120.

Carey, B., O'Brian, S., Onslow, M., Packman, A. & Menzies, R. (2012) Webcam delivery of the Camperdown Program for adolescents who stutter: A Phase I trial. *Language, Speech, and Hearing Services in Schools*. 43 (3), 370–380.

Cavenagh, P., Costelloe, S., Davis, S. & Howell, P. (2015) Characteristics of young children close to the onset of stuttering. *Communication Disorders Quarterly*. 36 (3), 162–171.

Central Department of Statistics and Information. (2011) *Latest Statistical Releases. Riyadh (KSA): Central Department of Statistics and Information, 2011*. Available from: http:// www.cdsi.gov.sa/english/ [Accessed 7th February 2016].

Chang, S.E., Zhu, D.C., Choo, A.L. & Angstadt, M. (2015) White matter neuroanatomical differences in young children who stutter. *Brain*. 138 (3), 694–711.

Chang, S.E., Erickson, K.I., Ambrose, N.G., Hasegawa-Johnson, M.A. & Ludlow, C.L. (2008) Brain anatomy differences in childhood stuttering. *Neuroimage*. 39 (3), 1333–1344.

Chon, H., Sawyer, J. & Ambrose, N.G. (2012) Differences of articulation rate and utterance length in fluent and disfluent utterances of preschool children who stutter. *Journal of Communication Disorders*. 45 (6), 455–467.

Choo, A.L., Kraft, S.J., Olivero, W., Ambrose, N.G., Sharma, H., Chang, S.E. & Loucks, T.M. (2011) Corpus callosum differences associated with persistent stuttering in adults. *Journal of Communication Disorders*. 44 (4), 470–477.

Ciccia, A.H., Whitford, B., Krumm, M. & McNeal, K. (2011) Improving the access of young urban children to speech, language and hearing screening via telehealth. *Journal of Telemedicine and Telecare*. 17 (5), 240–244.

Clark, C.E., Conture, E.G., Frankel, C.B. & Walden, T.A. (2012) Communicative and psychological dimensions of the KiddyCAT. *Journal of Communication Disorders*. 45 (3), 223–234.

Cocomazzo, N., Block, S., Carey, B., O'Brian, S., Onslow, M., Packman, A. & Iverach, L. (2012) Camperdown Program for adults who stutter: student training clinic Phase I trial. *International Journal of Language & Communication Disorders*. 47 (4), 365–372.

Cohn, E.R. & Watzlaf, V.J. (2011) Privacy and internet-based telepractice. *Perspectives on Telepractice*. 1 (1), 26–27.

Coleman, J. (1999) *The Nature of Adolescence*. Routledge.

Cooper, E.B. & Cooper, C. (1985) *Personalized fluency control therapy*. DLM Teaching Resources.

Communications and Information Technology Commission (CITC). (2014) *Annual report 2014*. Available from: http://www.citc.gov.sa/english/Pages/default.aspx [Accessed 7th February 2016].

Communications and Information Technology Commission (CITC). (2012) Available from: http://www.citc.gov.sa/en/MediaCenter/CITCinthemedia/Pages/PR_MED_127.aspx [Accessed 7th February 2016].

Constantino, C.D., Leslie, P., Quesal, R.W. & Yaruss, J.S. (2016) A preliminary investigation of daily variability of stuttering in adults. *Journal of Communication Disorders*. 60, 39–50.

Constantinescu, G., Theodoros, D., Russell, T., Ward, E., Wilson, S. & Wootton, R. (2010) Assessing disordered speech and voice in Parkinson's disease: A telerehabilitation application. International Journal of Language & Communication Disorders. 45 (6), 630–644.

Constantinescu, G., Theodoros, D., Russell, T., Ward, E., Wilson, S. & Wootton, R. (2011) Treating disordered speech and voice in Parkinson's disease online: A randomized controlled non- inferiority trial. International Journal of Language & Communication Disorders. 46 (1), 1–16.

Cook, F., Botterill, W., Lees, R. & Stark, C. (2005) Family-based approach to therapy with primary school children: 'Throwing the ball back'. In: Lees, R. & Stark, C. (eds.) The treatment of stuttering in the young school-aged child. London: Whurr, pp.81–107.

Cook, S., Rieger, M., Donlan, C. & Howell, P. (2011) Testing orofacial abilities of children who stutter: The Movement, Articulation, Mandibular and Sensory awareness (MAMS) assessment procedure. Journal of Fluency Disorders, 36 (1), 27–40.

Cooper, E.B. (1966) Client-clinician relationships and concomitant factors in stuttering therapy. Journal of Speech, Language, and Hearing Research. 9 (2), 194–207.

Cordes, A.K. & Ingham, R.J. (1994) The reliability of observational data: II. Issues in the identification and measurement of stuttering events. Journal of Speech, Language, and Hearing Research. 37 (2), 279–294.

Cordes, A.K. & Ingham, R.J. (1995) Judgments of stuttered and non-stuttered intervals by recognized authorities in stuttering research. Journal of Speech, Language, and Hearing Research. 38 (1), 33–41.

Cordes, A.K. & Ingham, R.J. (1999) Effects of time-interval judgment training on real-time measurement of stuttering. Journal of Speech, Language, and Hearing Research. 42 (4), 862–879.

Cordes, A.K., Ingham, R.J., Frank, P. & Ingham, J.C. (1992) Time-interval analysis of interjudge and intrajudge agreement for stuttering event judgments. Journal of Speech, Language, and Hearing Research. 35 (3), 483–494.

Costello, J.M., Ingham, R.J., Curlee, R.F. & Perkins, W.H. (1984) Assessment strategies for stuttering. Nature and Treatment of Stuttering: New Directions. 303–333.

Craig, A.R. & Calver, P. (1991) Following up on treated stutterers: Studies of perceptions of fluency and job status. Journal of Speech, Language, and Hearing Research. 34 (2), 279–284.

Craig, A., Hancock, K., Chang, E., McCready, C., Shepley, A., McCaul, A., Costello, D., Harding, S., Kehren, R., Masel, C. & Reilly, K. (1996) A controlled clinical trial for stuttering in persons aged 9 to 14 years. Journal of Speech, Language, and Hearing Research. 39(4), 808-826.

Craig, A., Hancock, K., Tran, Y., Craig, M. & Peters, K. (2002) Epidemiology of stuttering in the community across the entire life span. *Journal of Speech, Language, and Hearing Research*. 45 (6), 1097–1105.

Craig, A. & Tran, Y. (2005) The epidemiology of stuttering: The need for reliable estimates of prevalence and anxiety levels over the lifespan. *Advances in Speech Language Pathology*. 7 (1), 41–46.

Crais, E.R., Roy, V.P. & Free, K. (2006) Parents' and professionals' perceptions of the implementation of family-centered practices in child assessments. *American Journal of Speech-Language Pathology*. 15 (4), 365–377.

Cream, A., Onslow, M., Packman, A. & Llewellyn, G. (2003) Protection from harm: The experience of adults after therapy with prolonged- speech. *International Journal of Language & Communication Disorders*. 38 (4), 379–395.

Crichton-Smith, I. (2002) Communicating in the real world: Accounts from people who stammer. *Journal of Fluency Disorders*, 27 (4), 333–352.

Crichton-Smith, I., Baker, K.L. & Rowley, D.T. (2003) Changing conversational dynamics: A case study in parent–child interaction therapy. In *Proceedings of the Sixth Oxford Dysfluency Conference*. pp.29–136.

Crutchley, S. & Campbell, M. (2010) TeleSpeech therapy pilot project: Stakeholder satisfaction. *International Journal of Telerehabilitation*. 2 (1), 23.

Cullinan, W.L., Prather, E.M. & Williams, D.E. (1963) Comparison of procedures for scaling severity of stuttering. *Journal of Speech, Language, and Hearing Research*. 6 (2), 187–194.

Curran, M.F. & Hood, S.B. (1977) Listener ratings of severity for specific disfluency types in children. *Journal of Fluency Disorders*. 2 (2), 87–97.

Davidow, J.H. & Scott, K.A. (2017) Intrajudge and Interjudge Reliability of the Stuttering Severity Instrument–Fourth Edition. *American Journal of Speech-Language Pathology*. 26 (4), 1105–1119.

Davis, S., Howell, P. & Cooke, F. (2002) Sociodynamic relationships between children who stutter and their non- stuttering classmates. *Journal of Child Psychology and Psychiatry*. 43 (7), 939–947.

Dawoud, A., Althbiti, A.A.J., Al Khatib, F.M. & AL-Ghalayini, N.A. (2017) Telemedicine: Between reality and challenges in Jeddah hospitals. *Egyptian Journal of Hospital Medicine*. 68 (3).

DePalma, J.A. (2009) Telehealth in the community: A research update. *Home Health Care Management & Practice*. 21 (3), 205–207.

De Nil, L.F. & Kroll, R.M. (2001) Searching for the neural basis of stuttering treatment outcome: recent neuroimaging studies. *Clinical Linguistics & Phonetics*. 15 (1-2), 163–168.

De Nil, L.F., Kroll, R.M., Lafaille, S.J. & Houle, S. (2003) A positron emission tomography study of short-and long-term treatment effects on functional brain activation in adults who stutter. *Journal of Fluency Disorders*. 28 (4), 357–380.

de Sonneville-Koedoot, C., Stolk, E., Rietveld, T. & Franken, M.C. (2015) Direct versus indirect treatment for preschool children who stutter: The RESTART randomized trial. *PloS One*. 10 (7), e0133758.

de Vet, H.C., Terwee, C.B., Knol, D.L. & Bouter, L.M. (2006) When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*. 59 (10), 1033–1039.

Domholdt, E. (2005) *Rehabilitation Research: Principles and Applications*. Elsevier Saunders.

Donaghy, M.A. & Smith, K.A. (2016) Management options for pediatric patients who stutter: Current challenges and future directions. *Pediatric Health, Medicine and Therapeutics*. 7, 71.

Druce, T., Debney, S. & Byrt, T. (1997) Evaluation of an intensive treatment program for stuttering in young children. *Journal of Fluency Disorders*. 28, 209 – 218.

Dudding, C.C. (2009) Digital videoconferencing: Applications across the disciplines. *Communication Disorders Quarterly*. 30 (3), 178–182.

Dworzynski, K., Remington, A., Rijsdijk, F., Howell, P. & Plomin, R. (2007) Genetic etiology in cases of recovered and persistent stuttering in an unselected, longitudinal sample of young twins. *American Journal of Speech-Language Pathology*. 16 (2), 169–178.

Edwards, M., Stredler-Brown, A. & Houston, K.T. (2012) Expanding use of telepractice in speech-language pathology and audiology. *Volta Review*, 112 (3).

El-Din, M.M.N., Al-Shakhs, F.N. & Al-Oudah, S.S. (2008) Missed appointments at a university hospital in eastern Saudi Arabia: Magnitude and association factors. *The Journal of the Egyptian Public Health Association*. 83 (5–6), 415–433.

Eriks- Brophy, A., Quittenbaum, J., Anderson, D. & Nelson, T. (2008) Part of the problem or part of the solution? Communication assessments of Aboriginal children residing in remote communities using videoconferencing. *Clinical Linguistics & Phonetics*, 22 (8), 589–609.

Erickson, S. & Block, S. (2013) The social and communication impact of stuttering on adolescents and their families. *Journal of Fluency Disorders*. 38 (4), 311–324.

Erickson, S., Block, S., Menzies, R., O'Brian, S., Packman, A. & Onslow, M. (2016) Standalone internet speech restructuring treatment for adults who stutter: A phase I study. *International Journal of Speech-Language Pathology*. 18 (4), 329–340.

Erickson, S., Block, S., Menzies, R., Onslow, M., O'Brian, S. & Packman, A. (2012) Standalone internet speech restructuring treatment for adults who stutter: A pilot study. *Journal of Clinical Practice in Speech-Language Pathology*. 14, 118–123.

Eve, C.L., Onslow, M., Andrews, C. & Adams, R. (1995) Clinical measurement of early stuttering severity: The reliability of a 10-point scale. *Australian Journal of Human Communication Disorders*. 23 (2), 26–39.

Ezrati-Vinacour, R. & Levin, I. (2004) The relationship between anxiety and stuttering: A multidimensional approach. *Journal of Fluency Disorders*, 29 (2), 135–148.

Fagan, J. & Jacobs, M. (2009) Survey of ENT services in Africa: Need for a comprehensive intervention. *Global Health Action*. 2 (1), 1932.

Fairweather, G.C., Lincoln, M.A. & Ramsden, R. (2016) Speech-language pathology teletherapy in rural and remote educational settings: Decreasing service inequities. *International journal of Speech-Language Pathology*. 18 (6), 592–602.

Fairweather, C., Parkin, M., Rozsa, M., Murdoch, B.E., Goozee, J., Whelan, B. & Docking, K. (2004) Speech and language assessment in school-aged children via videoconferencing. In: *Proceedings of the 26th World Congress of the International Association of Logopaedics and Phoniatrics. Melbourne, Speech Pathology Australia*.

Felsenfeld, S., Kirk, K.M., Zhu, G., Statham, D.J., Neale, M.C. & Martin, N.G. (2000) A study of the genetic and environmental etiology of stuttering in a selected twin sample. *Behavior Genetics*. 30 (5), 359–366.

Fibiger, S., Frøkiær, L. & Jensen, T.M. (2007) Visual analog scale self-evaluation of stuttering aspects before and after therapy in secondary school children—A pilot study. In: *27th World Congress of the International Association of Logopedics and Phoniatrics, Copenhagen, August*.

Fischer, A.J., Dart, E.H., Leblanc, H., Hartman, K.L., Steeves, R.O. & Gresham, F.M. (2016) An investigation of the acceptability of videoconferencing within a school-based behavioural consultation framework. *Psychology in the Schools*. 53 (3), 240–252.

Fleming, A.M., Brown, K.J. & Houston, K.T. (2013) Putting the 'tele-' in telepractice. *SIG 18 Perspectives on Telepractice*. 3, 9–15.

Foundas, A.L., Bollich, A.M., Corey, D.M., Hurley, M. & Heilman, K.M. (2001) Anomalous anatomy of speech–language areas in adults with persistent developmental stuttering. *Neurology*. 57 (2), 207–215.

Fox, P.T., Ingham, R.J., Ingham, J.C., Zamarripa, F., Xiong, J.H. & Lancaster, J.L. (2000) Brain correlates of stuttering and syllable production: A PET performance-correlation analysis. *Brain*. 123 (10), 1985– 2004.

Franic, D.M. & Bothe, A.K. (2008) Psychometric evaluation of condition-specific instruments used to assess health-related quality of life, attitudes, and related constructs in stuttering. *American Journal of Speech-Language Pathology*. 17 (1), 60–80.

Franken, M.C.J., Kielstra-Van der Schalk, C.J. & Boelens, H. (2005) Experimental treatment of early stuttering: A preliminary study. *Journal of Fluency Disorders*. 30 (3), 189–199.

Galvan, C., Case, E. & Houston, K.T. (2014) Listening and learning: Using telepractice to serve children and adults with hearing loss. *SIG 18 Perspectives on Telepractice*. 4 (1), 11–22.

Giavarina, D. (2015) Understanding Bland Altman analysis. *Biochemia Medica.* 25 (2), 141–151.

Goldberg, S., Haley, K.L. & Jacks, A. (2012) Script training and generalization for people with aphasia. *American Journal of Speech-Language Pathology*. 21 (3), 222–238.

Griffiths J. (2001) Missed appointments rise by 60%. *GP 2001*. August 24: 7.

Grogan-Johnson, S. (2012) Providing school-based speech-language therapy services by telepractice: A brief tutorial. *SIG 18 Perspectives on Telepractice* 2 (1), 42–48.

Grogan-Johnson, S., Alvares, R., Rowan, L. & Creaghead, N. (2010) A pilot study comparing the effectiveness of speech language therapy provided by telemedicine with conventional on-site therapy. *Journal of Telemedicine and Telecare*. 16 (3), 134–139.

Grogan-Johnson, S., Gabel, R.M., Taylor, J., Rowan, L.E., Alvares, R. & Schenker, J. (2011) A pilot exploration of speech sound disorder intervention delivered by telehealth to school–age children. *International Journal of Telerehabilitation*. 3 (1), 31.

Grogan-Johnson, S., Schmidt, A.M., Schenker, J., Alvares, R., Rowan, L.E. & Taylor, J. (2013) A comparison of speech sound intervention delivered by telepractice and side-by-side service delivery models. *Communication Disorders Quarterly*. 34 (4), 210–220.

Guitar, B. (2006) *Stuttering: An Integrated Approach to Its Nature and Treatment*. Lippincott Williams & Wilkins.

Guntupalli, V.K., Kalinowski, J. & Saltuklaroglu, T. (2006) The need for self- report data in the assessment of stuttering therapy efficacy: Repetitions and prolongations of speech. The stuttering syndrome. *International Journal of Language & Communication Disorders*. 41 (1), 1–18.

Hall, K.D., Amir, O. & Yairi, E. (1999) A longitudinal investigation of speaking rate in preschool children who stutter. *Journal of Speech, Language, and Hearing Research*. 42 (6), 1367–1377.

Hall, D.E., Lynn, J.M., Altieri, J., Segers, V.D. & Conti, D. (1987) Inter-intrajudge reliability of the stuttering severity instrument. *Journal of Fluency Disorders*. 12 (3), 167–173.

Hallgren, K.A. (2012) Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology*. 8 (1), 23.

Hancock, K., Craig, A., McCready, C., McCaul, A., Costello, D., Campbell, K. & Gilmore, G. (1998) Two-to six-year controlled-trial stuttering outcomes for children and adolescents. *Journal of Speech, Language, and Hearing Research*. 41(6), 1242-1252.

Hargrave, S., Kalinowski, J., Stuart, A., Armson, J. & Jones, K. (1994) Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rates. *Journal of Speech, Language, and Hearing Research*. 37 (6), 1313–1319.

Harris, V., Onslow, M., Packman, A., Harrison, E. & Menzies, R. (2002) An experimental investigation of the impact of the Lidcombe Program on early stuttering. *Journal of Fluency Disorders*. 27 (3), 203–214.

Harrison, E., Wilson, L. & Onslow, M. (1999) Distance intervention for early stuttering with the Lidcombe Programme. *Advances in Speech Language Pathology*. 1 (1), 31–36.

Hayhow, R., Cray, A.M. & Enderby, P. (2002) Stammering and therapy views of people who stammer. *Journal of Fluency Disorders*, 27 (1), 1–17.

Healey, E.C. (1991) Stuttering severity instrument for children and adults: (revised). Riley, G. (1986). Austin TX: Pro-Ed. *Journal of Fluency Disorders*. 16 (5–6), 309–311.

Hearne, A., Packman, A., Onslow, M. & O'Brian, S. (2008) Developing treatment for adolescents who stutter: A Phase I trial of the Camperdown Program. *Language, Speech, and Hearing Services in Schools*. 39 (4), 487–497.

Hill, A. & Theodoros, D. (2002) Research into telehealth applications in speech-language pathology. *Journal of Telemedicine and Telecare*. 8 (4), 187–196.

Hill, A.J., Theodoros, D.G., Russell, T.G., Cahill, L.M., Ward, E.C. & Clark, K.M. (2006) An internet-based telerehabilitation system for the assessment of motor speech disorders: A pilot study. *American Journal of Speech-Language Pathology*. 15 (1), 45–56.

Hill, A.J., Theodoros, D., Russell, T. & Ward, E. (2009) Using telerehabilitation to assess apraxia of speech in adults. *International Journal of Language & Communication Disorders*. 44 (5), 731–747.

Hill, A.J., Theodoros, D.G., Russell, T.G. & Ward, E.C. (2009a) The redesign and re-evaluation of an internet-based telerehabilitation system for the assessment of dysarthria in adults. *Telemedicine and e-Health*. 15 (9), 840–850.

Hill, A.J., Theodoros, D.G., Russell, T.G., Ward, E.C. & Wootton, R. (2009b) The effects of aphasia severity on the ability to assess language disorders via telerehabilitation. *Aphasiology*. 23 (5), 627–642.

Hines, M., Lincoln, M., Ramsden, R., Martinovich, J. & Fairweather, C. (2015) Speech pathologists' perspectives on transitioning to telepractice: What factors promote acceptance? *Journal of Telemedicine and Telecare*. 21 (8), 469–473.

Houston, K.T. (2013) *Telepractice in Speech-Language Pathology*. Plural Publishing.

Howell, P. (2011) *Recovery from Stuttering*. Routledge.

Howell, P. & Davis, S. (2011) Predicting persistence of and recovery from stuttering by the teenage years based on information gathered at age 8 years. *Journal of Developmental & Behavioral Pediatrics*. 32 (3), 196–205.

Howell, P., Davis, S. & Williams, R. (2008) Late childhood stuttering. *Journal of Speech, Language, and Hearing Research*. 51 (3), 669–687.

Howie, P.M. (1981) Concordance for stuttering in monozygotic and dizygotic twin pairs. *Journal of Speech, Language, and Hearing Research*. 24 (3), 317–321.

Howie, P.M., Tanner, S. & Andrews, G. (1981) Short-and long-term outcome in an intensive treatment program for adult stutterers. *Journal of Speech and Hearing Disorders*. 46(1),104-109.

Iacono, T., Johnson, H., Humphreys, J. & McAllister, L. (2007) Recruitment of speech pathologists into positions considered less attractive. *Advances in Speech Language Pathology*. 9 (3), 204–212.

Ingham, R.J. & Cordes, A.K. (1992) Interclinic differences in stuttering-event counts. *Journal of Fluency Disorders*. 17 (3), 171–176.

Ingham, R.J., Gow, M. & Costello, J.M. (1985) Stuttering and speech naturalness: Some additional data. *Journal of Speech and Hearing Disorders.* 50 (2), 217–219.

Ingham, R.J., Kilgo, M., Ingham, J.C., Moglia, R., Belknap, H. & Sanchez, T. (2001) Evaluation of a stuttering treatment based on reduction of short phonation intervals. *Journal of Speech, Language, and Hearing Research*. 44 (6), 1229–1244.

Ingham, R.J. & Onslow, M. (1985) Measurement and modification of speech naturalness during stuttering therapy. *Journal of Speech and Hearing Disorders*. 50 (3), 261–281.

Ingham, R.J. & Packman, A. (1977) Treatment and generalization effects in an experimental treatment for a stutterer using contingency management and speech rate control. *Journal of Speech and Hearing Disorders*. 42 (3), 394–407.

Irani, F. & Gabel, R. (2011) Intensive stuttering therapy with telepractice follow-up: A case study. *Perspectives on Fluency and Fluency Disorders.* 21 (1), 11–21.

Iverach, L., O'Brian, S., Jones, M., Block, S., Lincoln, M., Harrison, E., Hewat, S., Menzies, R.G., Packman, A. & Onslow, M. (2009) Prevalence of anxiety disorders among adults seeking speech therapy for stuttering. *Journal of Anxiety Disorders*. 23 (7), 928–934.

Jäncke, L., Hänggi, J. & Steinmetz, H. (2004) Morphological brain differences between adult stutterers and non-stutterers. *BMC Neurology*. 4 (1), 23.

Jani, L., Huckvale, M. & Howell, P. (2013) Procedures used for assessment of stuttering frequency and stuttering duration. *Clinical Linguistics & Phonetics*. 27 (12), 853–861.

Jiang, J., Lu, C., Peng, D., Zhu, C. & Howell, P. (2012) Classification of types of stuttering symptoms based on brain activity. *PloS One*. 7 (6), e39747.

Johannisson, T.B., Wennerfeldt, S., Havstam, C., Naeslund, M., Jacobson, K. & Lohmander, A. (2009) The Communication Attitude Test (CAT- S): Normative values for 220 Swedish children. *International Journal of Language & Communication Disorders*. 44 (6), 813–825.

Jones, M., Dobson, A. & O'Brian, S. (2011) A graphical method for assessing agreement with the mean between multiple observers using continuous measures. *International Journal of Epidemiology*. 40 (5), 1308–1313.

Jones, M., Gebski, V., Onslow, M. & Packman, A. (2002) Statistical power in stuttering research: A tutorial. *Journal of Speech, Language, and Hearing Research*. 45 (2), 243–255.

Jones, M., Onslow, M., Packman, A., Williams, S., Ormond, T., Schwarz, I. & Gebski, V. (2005) Randomised controlled trial of the Lidcombe Program of early stuttering intervention. *Bmj*. 331 (7518), 659.

Juenger, J.M. (2009) Telepractice in the schools: Virtual services help clinicians do more with less. *The ASHA Leader*. 14 (12), 20–21.

Kang, C., Riazuddin, S., Mundorff, J., Krasnewich, D., Friedman, P., Mullikin, J.C. & Drayna, D. (2010) Mutations in the lysosomal enzyme–targeting pathway and persistent stuttering. *New England Journal of Medicine*. 362 (8), 677–685.

Karimi, H., Jones, M., O'Brian, S. & Onslow, M. (2014) Clinician percent syllables stuttered, clinician severity ratings and speaker severity ratings: Are they interchangeable? *International Journal of Language & Communication Disorders*. 49 (3), 364–368.

Karimi, H., O'Brian, S., Onslow, M. & Jones, M. (2014a) Absolute and relative reliability of percentage of syllables stuttered and severity rating scales. *Journal of Speech, Language, and Hearing Research*. 57 (4), 1284–1295.

Karimi, H., O'Brian, S., Onslow, M., Jones, M., Menzies, R. & Packman, A. (2013) Using statistical process control charts to study stuttering frequency variability during a single day. *Journal of Speech, Language, and Hearing Research*. 56 (6), 1789–1799.

Keck, C.S. & Doarn, C.R. (2014) Telehealth technology applications in speech-language pathology. *Telemedicine and e-Health*. 20 (7), 653–659.

Kefalianos, E., Onslow, M., Packman, A., Vogel, A., Pezic, A., Mensah, F., Conway, L., Bavin, E., Block, S. & Reilly, S. (2017) The history of stuttering by 7 years of age: Follow-up of a prospective community cohort. *Journal of Speech, Language, and Hearing Research*. 60 (10), 2828–2839.

Kell, C.A., Neumann, K., von Kriegstein, K., Posenenske, C., von Gudenberg, A.W., Euler, H. & Giraud, A.L. (2009) How the brain repairs stuttering. *Brain*. 132 (10), 2747–2760.

Kelman, E. & Nicholas, A. (2017) *Practical Intervention For Early Childhood Stammering: Palin PCI Approach*. Routledge.

Kelly, E.M., Martin, J.S., Baker, K.E., Rivera, N.I., Bishop, J.E., Krizizke, C.B., Stettler, D.S. & Stealy, J.M. (1997) Academic and clinical preparation and practices of school speech-language pathologists with people who stutter. *Language, Speech, and Hearing Services in Schools*. 28 (3), 195–212.

Khoja, M.A. (2017) A survey of formal and informal assessment procedures used by speech-language pathologists in Saudi Arabia. *Speech, Language and Hearing*. 1–9.

Khoja, M.A. & Sheeshah, H. (2018) The human right to communicate: A survey of available services in Saudi Arabia. *International Journal of Speech-Language Pathology*. 20 (1), 102–107.

Kidd, K.K., Heimbuch, R.C. & Records, M.A. (1981) Vertical transmission of susceptibility to stuttering with sex-modified expression. *Proceedings of the National Academy of Sciences*. 78 (1), 606–610.

Klein, J.F. & Hood, S.B. (2004) The impact of stuttering on employment opportunities and job performance. *Journal of Fluency Disorders*, 29 (4), 255–273.

Koch, S. (2006) Home telehealth—current state and future trends. *International Journal of Medical Informatics*. 75 (8), 565–576.

Koo, T.K. & Li, M.Y. (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*. 15 (2), 155–163.

Kraaimaat, F.W., Vanryckeghem, M. & Van Dam-Baggen, R. (2002) Stuttering and social

anxiety. *Journal of Fluency Disorders*. 27 (4), 319–331.

Kraft, S.J. & Yairi, E. (2012) Genetic bases of stuttering: The state of the art, 2011. *Folia Phoniatrica et Logopaedica*. 64 (1), 34–47.

Krupinski, E.A., Charness, N., Demiris, G., Bulik, R.J. & Seale, D.E. (2008) Human factors in telemedicine. *Telemedicine and e-health*. 14 (10), 1024–1030.

Kully, D. (2000) Telehealth in speech pathology: Applications to the treatment of stuttering. *Journal of Telemedicine and Telecare*. 6 (2_suppl), 39–41.

Kully, D. (2002) Venturing into telehealth: Applying interactive technologies to stuttering treatment. *The ASHA Leader*. 7 (11), 1–15.

Kully, D. & Boberg, E. (1988) An investigation of interclinic agreement in the identification of fluent and stuttered syllables. *Journal of Fluency Disorders*. 13 (5), 309–318.

Lan, J., Song, M., Pan, C., Zhuang, G., Wang, Y., Ma, W., Chu, Q., Lai, Q., Xu, F., Li, Y. & Liu, L. (2009) Association between dopaminergic genes (SLC6A3 and DRD2) and stuttering among Han Chinese. *Journal of Human Genetics*. 54 (8), 457.

Landis, J.R. & Koch, G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics*. 159–174.

Langevin, M., Packman, A. & Onslow, M. (2010) Parent perceptions of the impact of stuttering on their preschoolers and themselves. *Journal of Communication Disorders*. 43 (5), 407–423.

Lasker, J.P., Stierwalt, A.J., Spence, M. & Cavin-Root, C. (2010) Using webcam interactive technology to implement treatment for severe apraxia: A case example. *Journal of Medical Speech-Language Pathology*. 18 (4), 71–77.

Lattermann, C., Euler, H.A. & Neumann, K. (2008) A randomized control trial to investigate the impact of the Lidcombe Program on early stuttering in German-speaking preschoolers. *Journal of Fluency Disorders*. 33 (1), 52–65.

Lattermann, C., Shenker, R.C. & Thordardottir, E. (2005) Progression of language complexity during treatment with the Lidcombe Program for early stuttering intervention. *American Journal of Speech-Language Pathology*. 14 (3), 242–253.

Lee, P., Li, P.C., Liu, C.H. & Hsieh, C.L. (2011) Test–retest reliability of two attention tests in schizophrenia. *Archives of Clinical Neuropsychology*. 26 (5), 405–411.

Lewis, K.E. (1995) Do SSI-3 scores adequately reflect observations of stuttering behaviors? *American Journal of Speech-Language Pathology*. 4 (4), 46–59.

Lewis, C., Packman, A., Onslow, M., Simpson, J.M. & Jones, M. (2008) A Phase II trial of telehealth delivery of the Lidcombe Program of Early Stuttering Intervention. *American Journal of Speech-Language Pathology*. 17 (2), 139–149.

Lewis, D. and Sherman, D. (1951) Measuring the severity of stuttering. *Journal of Speech & Hearing Disorders*. 16, 320-326.

Liaw, L.J., Hsieh, C.L., Lo, S.K., Chen, H.M., Lee, S. & Lin, J.H. (2008) The relative and absolute reliability of two balance performance measures in chronic stroke patients. *Disability and Rehabilitation*. 30 (9), 656–661.

Lincoln, M.A. & Onslow, M. (1997) Long-term outcome of early intervention for stuttering. *American Journal of Speech-Language Pathology*. 6 (1), 51–58.

Logan, K.J. & Haj-Tas, M.A. (2007) Effect of sample size on the measurement of stutter-like disfluencies. *Perspectives on Fluency and Fluency Disorders*. 17 (3), 3–6.

LoPresti, E.F., Jinks, A. & Simpson, R.C., 2015. Consumer satisfaction with telerehabilitation service provision of alternative computer access and augmentative and alternative communication. *International Journal of Telerehabilitation*. 7 (2), 3.

Lowe, R., O'Brian, S. & Onslow, M. (2013) Review of telehealth stuttering management. *Folia Phoniatrica et Logopaedica*. 65 (5), 223–238.

Lu, C., Peng, D., Chen, C., Ning, N., Ding, G., Li, K., Yang, Y. & Lin, C. (2010) Altered effective connectivity and anomalous anatomy in the basal ganglia-thalamocortical circuit of stuttering speakers. *Cortex*. 46 (1), 49–67.

Luc, F. & Brutten, G.J. (1990) Speech-associated attitudes: Stuttering, voice disordered, articulation disordered, and normal speaking children. *Journal of Fluency Disorders*. 15 (2), 127–134.

Mansouri-Rad, P., Mahmood, A.M., Thompson, S.E. & Putnam, K. (2013) Culture matters: Factors affecting the adoption of telemedicine. In: *2013 46th Hawaii International Conference on System Sciences*. IEEE. pp. 2515–2524.

Månsson, H. (2000) Childhood stuttering: Incidence and development. *Journal of Fluency Disorders*. 25 (1), 47–57.

Martin, R.R., Haroldson, S.K. & Triden, K.A. (1984) Stuttering and speech naturalness. *Journal of Speech and Hearing Disorders*. 49 (1), 53–58.

Mashima, P.A., Birkmire-Peters, D.P., Syms, M.J., Holtel, M.R., Burgess, L.P. & Peters, L.J. (2003) Telehealth: Voice therapy using telecommunications technology. *American Journal of Speech-Language Pathology*. 12 (4), 432–439.

Mashima, P.A. & Doarn, C.R. (2008) Overview of telehealth activities in speech-language pathology. *Telemedicine and e-Health*. 14 (10), 1101–1117.

Matthews, R.A. (2014) The 'acceptability' of Skype mediated Speech and Language Therapy provision to school aged language impaired children (Doctoral dissertation, UCL [University College London]).

Matthews, S., Williams, R. & Pring, T. (1997) Parent—child interaction therapy and dysfluency: A single- case study. International Journal of Language & Communication Disorders. 32 (3), 346–357.

May, J. & Erickson, S. (2014) Telehealth: Why not? Journal of Clinical Practice in Speech-Language Pathology. 16 (3), 147–151.

McAllister, J., Collier, J. & Shepstone, L. (2013) The impact of adolescent stuttering and other speech problems on psychological well- being in adulthood: evidence from a birth cohort study. International Journal of Language & Communication Disorders. 48 (4), 458–468.

McCauley, R. (1989) Measurement as a dangerous activity. Journal of Speech-Language Pathology and Audiology. 13, 29–32.

McCue, M., Fairman, A. & Pramuka, M. (2010) Enhancing quality of life through telerehabilitation. Physical Medicine and Rehabilitation Clinics. 21 (1), 195–205.

McGuire, D. (2003) Beyond stuttering. The McGuire programme for getting good at the sport of speaking. Souvenir Press.

McKinnon, D.H., McLeod, S. & Reilly, S. (2007) The prevalence of stuttering, voice, and speech-sound disorders in primary school students in Australia. Language, Speech, and Hearing Services in Schools. 38 (1), 5–15.

McLeod, S. & Harrison, L.J. (2009) Epidemiology of speech and language impairment in a nationally representative sample of 4-to 5-year-old children. Journal of Speech, Language, and Hearing Research. 52 (5), 1213–1229.

Menzies, R.G., O'Brian, S., Onslow, M., Packman, A., St Clare, T. & Block, S. (2008) An experimental clinical trial of a cognitive-behavior therapy package for chronic stuttering. Journal of Speech, Language, and Hearing Research. 51 (6), 1451–1464.

Millard, S.K. & Davis, S. (2016) The Palin Parent Rating Scales: Parents' perspectives of childhood stuttering and its impact. Journal of Speech, Language, and Hearing Research. 59 (5), 950–963.

Millard, S.K., Edwards, S. & Cook, F.M. (2009) Parent-child interaction therapy: Adding to the evidence. International Journal of Speech-Language Pathology. 11 (1), 61–76.

Millard, S.K., Nicholas, A. & Cook, F.M. (2008) Is parent–child interaction therapy effective in reducing stuttering? *Journal of Speech, Language, and Hearing Research.* 51 (3), 636–650.

Miller, N., Noble, E., Jones, D. & Burn, D. (2006) Life with communication changes in Parkinson's disease. *Age and Ageing*. 35 (3), 235–239.

Ministry of Communications and Information Technology (MOCIT). (2003) *Information and telecommunication technology in Saudi Arabia (Report No. WSIS/PC- 3/CONTR/25-E)*. World Summit on the Information Society, Geneva.

Ministry of Health. (2015) *Health statistics yearbook, 2015*. Available from: http://www.moh.gov.sa/en/ministry/statistics/ book/pages/default.aspx [Accessed 3rd January 2017].

Ministry of Health. (2018) *MoH: All the Kingdom will be covered by telemedicine in two months*. Available from: https://www.moh.gov.sa/en/Ministry/MediaCenter/News/Pages/news-2018-03-06-006.aspx [Accessed 10th August 2018].

Ministry of Planning. (2000) *Demographic Features of KSA Populations:* Riyadh.

Miniwatts Marketing Group. (2013) *Internet world stats: Usage and population statistics*. Available from: http://www.internetworldstats.com/stats.htm [Accessed December 2014].

Mohamed, B.A. & Al-Doghaither, A.H. (2002) Missed appointments at public hospitals in Riyadh, Saudi Arabia. *Saudi Medical Journal*. 23 (4), 388–392.

Müller, R. & Büttner, P. (1994) A critical discussion of intraclass correlation coefficients. *Statistics in Medicine*. 13 (23–24), 2465–2476.

Murphy, A.T. & FitzSimons, R.M. (1960) *Stuttering and Personality Dynamics: Play Therapy, Projective Therapy*. Ronald Press Company.

Musa Jaber, M., Ghani, A., Khanapi, M. & Suryana Herman, N. (2014) A review of adoption of telemedicine in Middle East countries: Toward building an Iraqi Telemedicine Framework. *Science International*. 26 (5).

Myles, P.S. & Cui, J. (2007) Using the Bland–Altman method to measure agreement with repeated measures. *Anesthesia and analgesia.* 111 (5). 1180–1192.

Nelson, E.L. & Palsbo, S. (2006) Challenges in telemedicine equivalence studies. *Evaluation and Program Planning*. 29 (4), 419–425.

Nippold, M.A., 2011. Stuttering in school-age children: A call for treatment research. *Language, Speech, and Hearing Services in Schools*. 42 (2), 99–101.

O'Brian, S., Carey, B., Onslow, M., Packman, A. & Cream, A. (2010) *The Camperdown Program for Stuttering: Treatment Manual*. Australian Stuttering Research Center.

O'Brian, S., Jones, M., Lincoln, M., Harrison, E., Packman, A., Menzies, R. & Onslow, M. (2013) Validity of real-time measures of stuttering frequency. *Speech, Language and Hearing*. 16 (2), 107–116.

O'Brian, S., Jones, M., Onslow, M., Packman, A., Menzies, R. & Lowe, R. (2015) Comparison of audio and audio-visual measures of adult stuttering: Implications for clinical trials. *International Journal of Speech-Language Pathology*. 17 (6), 589–593.

O'Brian, S., Jones, M., Packman, A., Menzies, R. & Onslow, M. (2011) Stuttering severity and educational attainment. *Journal of Fluency Disorders*. 36 (2), 86–92.

O'Brian, S., Jones, M., Pilowsky, R., Onslow, M., Packman, A. & Menzies, R. (2010) A new method to sample stuttering in preschool children. *International Journal of Speech-Language Pathology*. 12 (3), 173–177.

O'Brian, S., Onslow, M., Cream, A. & Packman, A. (2003) The Camperdown Program: Outcomes of a new prolonged-speech treatment model. *Journal of Speech, Language, and Hearing Research*. 46 (4), 933–946.

O'Brian, S., Packman, A. & Onslow, M. (2004) Self-rating of stuttering severity as a clinical tool. *American Journal of Speech-Language Pathology*. 13 (3), 219–226.

O'Brian, S., Packman, A. & Onslow, M. (2008) Telehealth delivery of the Camperdown Program for adults who stutter: A phase I trial. *Journal of Speech, Language, and Hearing Research*. 51 (1), 184–195.

O'Brian, S., Packman, A., Onslow, M. & O'Brian, N. (2004) Measurement of stuttering in adults: Comparison of stuttering-rate and severity-scaling methods. *Journal of Speech, Language, and Hearing Research*. 47 (5), 1081–1087.

O'Brian, S., Smith, K. & Onslow, M. (2014) Webcam delivery of the Lidcombe Program for early stuttering: A phase I clinical trial. *Journal of Speech, Language, and Hearing Research*. 57 (3), 825–830.

O'Callaghan, A.M., McAllister, L. & Wilson, L. (2005) Barriers to accessing rural paediatric speech pathology services: Health care consumers' perspectives. *Australian Journal of Rural Health*. 13, 162–171.

Okalidou, A. & Kampanaros, M. (2001) Teacher perceptions of communication impairment at screening stage in preschool children living in Patras, Greece. *International Journal of Language & Communication Disorders*. 36 (4), 489–502.

Okasha, A., Bishry, Z., Kamel, M. & Hassan, A.H. (1974) Psychosocial study of stammering in Egyptian children. *The British Journal of Psychiatry*. 124 (583), 531–533.

Onslow, M. (2017) *Stuttering and Its Treatment: Eleven Lectures*.

Onslow, M., Adams, R. & Ingham, R. (1992) Reliability of speech naturalness ratings of stuttered speech during treatment. *Journal of Speech, Language, and Hearing Research*. 35 (5), 994–1001.

Onslow, M., Andrews, C. & Costa, L. (1990) Parental severity scaling of early stuttered speech: Four case studies. *Australian Journal of Human Communication Disorders*. 18 (1), 47–61.

Onslow, M., Andrews, C. & Lincoln, M. (1994) A control/experimental trial of an operant treatment for early stuttering. *Journal of Speech, Language, and Hearing Research*. 37 (6), 1244–1259.

Onslow, M., Costa, L., Andrews, C., Harrison, E. & Packman, A. (1996) Speech outcomes of a prolonged-speech treatment for stuttering. *Journal of Speech, Language, and Hearing Research*. 39 (4), 734–749.

Onslow, M., Jones, M., O'Brian, S., Packman, A. & Menzies, R. (2012) Stuttering. In: Sturmey, P. & Hersen, M. (eds.) *Handbook of Evidence-Based Practice in Clinical Psychology*. Hoboken, NJ: Wiley

Onslow, M., Jones, M., O'Brian, S., Packman, A., Menzies, R., Lowe, R., Arnott, S., Bridgman, K., de Sonneville, C. & Franken, M.C. (2018) Comparison of percentage of syllables stuttered with parent-reported severity ratings as a primary outcome measure in clinical trials of early stuttering treatment. *Journal of Speech, Language, and Hearing Research*. 61 (4), 811–819.

Onslow, M. & Millard, S. (2012) Palin Parent Child Interaction and the Lidcombe Program: Clarifying some issues. *Journal of Fluency Disorders*. 37 (1), 1–8.

Onslow, M., Packman, A. & Harrison, R.E. (2003) *The Lidcombe Program of Early Stuttering Intervention: A Clinician's Guide.* Austin, TX: Pro-Ed.

Onslow, M., Packman, A., Stocker, S., van Doorn, J. & Siegel, G.M. (1997) Control of children's stuttering with response-contingent time-out: Behavioral, perceptual, and acoustic data. *Journal of Speech, Language, and Hearing Research*. 40 (1), 121–133.

Packman, A. (2012) Theory and therapy in stuttering: A complex relationship. *Journal of Fluency Disorders*, 37 (4), 225–233.

Packman, A. & Attanasio, J.S. (2017) *Theoretical Issues in Stuttering*. Taylor & Francis.

Packman, A., Code, C. & Onslow, M. (2007) On the cause of stuttering: Integrating theory with brain and behavioral research. *Journal of Neurolinguistics*. 20 (5), 353–362.

Packman, A. & Meredith, G. (2011) Technology and the evolution of clinical methods for stuttering. *Journal of Fluency Disorders*. 36 (2), 75–85.

Packman, A., Onslow, M., Webber, M., Harrison, E., Arnott, S., Bridgeman, K., Carey, B., Sheedy, S., O'Brian, S., MacMillan, V. & Lloyd, W. (2014) *The Lidcombe Program of early stuttering intervention treatment manual*. Available from: http://sydney.edu.au/healthsciences/ asrc/health_ professionals/asrc_download. shtml. [Accessed 18th June 2017].

Palasik, S., Gabel, R., Hughes, C. & Rusnak, E. (2012) Perceptions about occupational experiences by people who stutter. *SIG 4 Perspectives on Fluency and Fluency Disorders*. 22 (1), 22–33.

Palsbo, S.E. (2007) Equivalence of functional communication assessment in speech pathology using videoconferencing. *Journal of Telemedicine and Telecare*. 13 (1), 40–43.

Perkins, W.H., Rudas, J., Johnson, L., Michael, W.B. & Curlee, R.F. (1974) Replacement of stuttering with normal speech: III. Clinical effectiveness. *Journal of Speech and Hearing Disorders*. 39 (4), 416–428.

Polovoy, C. (2008) Telepractice in schools helps address personnel shortages. *The ASHA Leader*. 13 (9), 22–24.

Proctor, A., Yairi, E., Duff, M.C. & Zhang, J. (2008) Prevalence of stuttering in African American preschoolers. *Journal of Speech, Language, and Hearing Research*. 51 (6), 1465–1479.

Rankin, G. & Stokes, M. (1998) Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clinical Rehabilitation*. 12 (3), 187–199.

Reilly, S. (2004) The challenges in making speech pathology practice evidence based. *Advances in Speech Language Pathology*. 6 (2), 113–124.

Reilly, S., Onslow, M., Packman, A., Cini, E., Conway, L., Ukoumunne, O.C., Bavin, E.L., Prior, M., Eadie, P., Block, S. & Wake, M. (2013) Natural history of stuttering to 4 years of age: A prospective community-based study. *Pediatrics*. 132 (3), 460–467.

Reilly, S., Onslow, M., Packman, A., Wake, M., Bavin, E.L., Prior, M., Eadie, P., Cini, E., Bolzonello, C. & Ukoumunne, O.C. (2009) Predicting stuttering onset by the age of 3 years: A prospective, community cohort study. *Pediatrics*. 123 (1), 270–277.

Reynolds, A.L., Vick, J.L. & Haak, N.J. (2009) Telehealth applications in speech-language pathology: A modified narrative review. *Journal of Telemedicine and Telecare*. 15 (6), 310–316.

Riaz, N., Steinberg, S., Ahmad, J., Pluzhnikov, A., Riazuddin, S., Cox, N.J. & Drayna, D. (2005) Genomewide significant linkage to stuttering on chromosome 12. *The American Journal of Human Genetics*. 76 (4), 647–651.

Riley, G. (1972) A stuttering severity instrument for children and adults. *Journal of Speech and Hearing Disorders*. 37 (3), 314–322.

Riley, J., Riley, G. & Maguire, G. (2004) Subjective screening of stuttering severity, locus of control and avoidance: Research edition. *Journal of Fluency Disorders*. 29 (1), 51–62.

Riley, G. (1994) *The Stuttering Severity Instrument for Adults and Children (SSI-3)*. 3rd ed. Austin, TX: PRO-ED.

Riley, G. (2009) *The Stuttering Severity Instrument for Adults and Children (SSI-4)*. 4th ed. Austin, TX: PRO-ED.

Roberts, P.M., Meltzer, A. & Wilding, J. (2009) Disfluencies in non-stuttering adults across sample lengths and topics. *Journal of Communication Disorders*. 42 (6), 414–427.

Rousseau, I., Onslow, M., Packman, A. & Jones, M. (2008) Comparisons of audio and audio-visual measures of stuttering frequency and severity in preschool-age children. *American Journal of Speech-Language Pathology*. 17 (2), 173–178.

Rousseau, I., Packman, A. & Onslow, M. (2005) *A trial of the Lidcombe Program with school age stuttering children*. In: Speech Pathology National Conference, Canberra, Australia.

Russell, T.G., Martin-Khan, M., Khan, A. & Wade, V. (2017) Method-comparison studies in telehealth: Study design and analysis considerations. *Journal of Telemedicine and Telecare*. 23 (9), 797–802.

Ryan, B.P. (2001) A longitudinal study of articulation, language, rate, and fluency of 22 preschool children who stutter. *Journal of Fluency Disorders*. 26 (2), 107–127.

Salmelin, R., Schnitzler, A., Schmitz, F., Jäncke, L., Witte, O.W. & Freund, H.J. (1998) Functional organization of the auditory cortex is different in stutterers and fluent speakers. *Neuroreport*. 9 (10), 2225–2229.

Salmelin, R., Schnitzler, A., Schmitz, F. & Freund, H.J. (2000) Single word reading in developmental stutterers and fluent speakers. *Brain*. 123 (6), 1184–1202.

Sawyer, J. & Yairi, E. (2006) The effect of sample size on the assessment of stuttering severity. *American Journal of Speech-Language Pathology*. 15 (1), 36–44.

Saudi Arabia Country Profile. (2010) Available from: http://www.emro.who.int [Accessed 1st April 2015].

Saudi Arabia Ministry of Health. (2010) *Population by Nationality in Health Regions 1430 H (2009 G.)*. Saudi Arabia Ministry of Health.

Sauerwein, L.B. & Linnemann, J.J. (2001) *Personal Data Protection Act*.

Schiavetti, N., Sacco, P.R., Metz, D.E. & Sitler, R.W. (1983) Direct magnitude estimation and interval scaling of stuttering severity. *Journal of Speech, Language, and Hearing Research*. 26 (4), 568–573.

Shaalan, S. (2009). Considerations for developing and adapting language and literacy assessments in Arabic-speaking countries. In: E. Grigorenko (ed.) *Multicultural psychoeducational assessment.* New York, Springer, pp. 287–314.

Sharma, S., Ward, E.C., Burns, C., Theodoros, D. & Russell, T. (2013) Assessing dysphagia via telerehabilitation: patient perceptions and satisfaction. *International Journal of Speech-Language Pathology*, 15 (2), 176–183.

Shaywitz, B.A., Shaywltz, S.E., Pugh, K.R., Constable, R.T., Skudlarski, P., Fulbright, R.K., Bronen, R.A., Fletcher, J.M., Shankweiler, D.P., Katz, L. & Gore, J.C. (1995) Sex differences in the functional organization of the brain for language. *Nature*. 373 (6515), 607.

Sherman, D. (1955) Reliability and utility of individual ratings of severity of audible characteristics of stuttering. *Journal of Speech and Hearing Disorders*. 20 (1), 11–16.

Shugart, Y.Y., Mundorff, J., Kilshaw, J., Doheny, K., Doan, B., Wanyee, J., Green, E.D. & Drayna, D. (2004) Results of a genome- wide linkage scan for stuttering. *American Journal of Medical Genetics Part A*. 124 (2), 133–135.

Sicotte, C., Lehoux, P., Fortier-Blanc, J. & Leblanc, Y. (2003) Feasibility and outcome evaluation of a telemedicine application in speech–language pathology. *Journal of Telemedicine and Telecare*. 9 (5), 253–258.

Sidavi, A. & Fabus, R. (2010) A review of stuttering intervention approaches for preschool-age and elementary school-age children. *Contemporary Issues in Communication Science Disorders*. *37*, 14–26.

Smith, A. & Kelly, E. (1997) Stuttering: A dynamic, multifactorial model. In: R. F. Curlee & G. M. Siegel (eds.) *Nature and treatment of stuttering: New directions*. 2nd ed. Boston, MA: Allyn & Bacon, pp. 204–217.

Smith, A. & Weber, C. (2016) Childhood stuttering–where are we and where are we going? *Seminars in Speech and Language*. 37 (4), 291.

Sommer, M., Koch, M.A., Paulus, W., Weiller, C. & Büchel, C. (2002) Disconnection of speech-relevant brain areas in persistent developmental stuttering. *The Lancet*. 360 (9330), 380–383.

Song, L.P., Peng, D.L., Jin, Z., Yao, L., Ning, N., Guo, X.J. & Zhang, T. (2007) Gray matter abnormalities in developmental stuttering determined with voxel-based morphometry. *Zhonghua Yi Xue Za Zhi*. 87 (41), 2884–2888.

Sønsterud, H., Mørk, M. & Lind, M. (2007) Improvement of fluency and listening skills in teenagers with advanced stuttering. In: *Proceedings of the 2007 IALP Congress.*

Stanberry, B. (2000) Telemedicine: barriers and opportunities in the 21st century. *Journal of Internal Medicine*, 247 (6), 615–628.

Starkweather, C.W. (1987) *Fluency and Stuttering*. Prentice-Hall.

Starkweather, C.W. & Gottwald, S.R. (1990) The demands and capacities model II: Clinical applications. *Journal of Fluency Disorders*. 15 (3), 143–157.

Stuart, A. & Kalinowski, J. (2004) The perception of speech naturalness of post-therapeutic and altered auditory feedback speech of adults with mild and severe stuttering. *Folia Phoniatrica et Logopaedica*. 56 (6), 347–357.

Suresh, R., Ambrose, N., Roe, C., Pluzhnikov, A., Wittke-Thompson, J.K., Ng, M.C.Y., Wu, X., Cook, E.H., Lundstrom, C., Garsten, M. & Ezrati, R. (2006) New complexities in the genetics of stuttering: significant sex-specific linkage signals. *The American Journal of Human Genetics*. 78 (4), 554–563.

Susca, M. (2006) Connecting stuttering measurement and management: II. Measures of cognition and affect. *International Journal of Language & Communication Disorders*. 41 (4), 365–377.

Sutherland, R., Trembath, D., Hodge, A., Drevensek, S., Lee, S., Silove, N. & Roberts, J. (2017) Telehealth language assessments using consumer grade equipment in rural and urban settings: Feasible, reliable and well tolerated. *Journal of Telemedicine and Telecare*. 23 (1), 106–115.

Tahmasebi, N., Shafie, B., Karimi, H. & Mazaheri, M. (2018) A Persian-version of the stuttering severity instrument-version four (SSI-4): How the new additions to SSI-4 complement its stuttering severity score? *Journal of Communication Disorders*. 74, 1–9.

Taylor, O.D., Armfield, N.R., Dodrill, P. & Smith, A.C. (2014) A review of the efficacy and effectiveness of using telehealth for paediatric speech and language assessment. *Journal of Telemedicine and Telecare*. 20 (7), 405–412.

Teesson, K., Packman, A. & Onslow, M. (2003) The Lidcombe behavioral data language of stuttering. *Journal of Speech, Language, and Hearing Research*. 46 (4), 1009–1015.

Tellis, G.M. (2007) Alarming statistics show training urgently needed in fluency disorders. *Stuttering Foundation of America Newsletter*. 5.

Teshima, S., Langevin, M., Hagler, P. & Kully, D. (2010) Post-treatment speech naturalness of comprehensive stuttering program clients and differences in ratings among listener groups. *Journal of Fluency Disorders*. 35 (1), 44–58.

Theodoros, D. (2012) A new era in speech-language pathology practice: Innovation and diversification. *International Journal of Speech-Language Pathology*. 14 (3), 189–199.

Theodoros, D. (2011) Telepractice in speech-language pathology: The evidence, the challenges, and the future. *Perspectives on Telepractice*. 1 (1), 10–21.

Theodoros, D. (2008) Telerehabilitation for service delivery in speech-language pathology. *Journal of Telemedicine and Telecare*. 14 (5), 221–224.

Theodoros, D., Hill, A., Russell, T., Ward, E. & Wootton, R. (2008) Assessing acquired language disorders in adults via the Internet. *Telemedicine and e-Health*. 14 (6), 552–559.

Theodoros, D., Russell, T.G., Hill, A., Cahill, L. & Clark, K. (2003) Assessment of motor speech disorders online: a pilot study. *Journal of Telemedicine and Telecare*, 9 (2_suppl), 66–68.

Throneburg, R.N. & Yairi, E. (2001) Durational, proportionate, and absolute frequency characteristics of disfluencies: A longitudinal study regarding persistence and recovery. *Journal of Speech, Language, and Hearing Research*. 44 (1), 38–51.

Todd, H., Mirawdeli, A., Costelloe, S., Cavenagh, P., Davis, S. & Howell, P. (2014) Scores on Riley's Stuttering Severity Instrument Versions Three and Four for samples of different

length and for different types of speech material. *Clinical Linguistics & Phonetics*. 28 (12), 912–926.

Tucker, J.K. (2012) Perspectives of speech-language pathologists on the use of telepractice in schools: Quantitative survey results. *International Journal of Telerehabilitation*. 4 (2), 61.

Tucker, J.K. (2012) Perspectives of speech-language pathologists on the use of telepractice in schools: The qualitative view. *International Journal of Telerehabilitation*. 4 (2), 47.

Tuthill, C. (1940) A quantitative study of extensional meaning with special reference to stuttering. *Journal of Speech Disorders*. 5 (2), 189–191.

UNICEF. (2013) *Saudi Arabia Statistics.* Available from: http://www.unicef.org/infobycountry/ saudiarabia_statistics.html [Accessed 14th June 2015].

Valente, A.R.S., Jesus, L.M., Hall, A. & Leahy, M. (2015) Event- and interval- based measurement of stuttering: A review. *International Journal of Language & Communication Disorders*. 50 (1), 14–30.

Valentine, D.T. (2014) Stuttering intervention in three service delivery models (direct, hybrid, and telepractice): Two case studies. *International Journal of Telerehabilitation*, 6 (2), 51.

Van Borsel, J. & Eeckhout, H. (2008) The speech naturalness of people who stutter speaking under delayed auditory feedback as perceived by different groups of listeners. *Journal of Fluency Disorders*. 33 (3), 241–251.

Van Borsel, J., Moeyaert, J., Mostaert, C., Rosseel, R., Van Loo, E. & Van Renterghem, T. (2006) Prevalence of stuttering in regular and special school populations in Belgium based on teacher perceptions. *Folia Phoniatrica et Logopaedica*. 58 (4), 289–302.

Van Eerdenbrugh, S., Packman, A., Onslow, M., O'Brian, S. & Menzies, R. (2018) Development of an internet version of the Lidcombe Program of early stuttering intervention: A trial of Part 1. *International Journal of Speech-Language Pathology*. 20 (2), 216–225.

Vangeneugden, T., Laenen, A., Geys, H., Renard, D. & Molenberghs, G. (2005) Applying concepts of generalizability theory on clinical trial data to investigate sources of variation and their impact on reliability. *Biometrics*. 61 (1), 295–304.

Van Riper, C. (1982) *The Nature of Stuttering*. Prentice Hall.

Van Stralen, K.J., Jager, K.J., Zoccali, C. & Dekker, F.W. (2008) Agreement between methods. *Kidney International*. 74 (9), 1116–1120.

Vanryckeghem, M. & Brutten, G. (2007) *The KiddyCAT: Communication Attitude Test for Preschool and Kindergarten Children Who Stutter.* Plural Publishing.

Vanryckeghem, M., Brutten, G.J. & Hernandez, L.M. (2005) A comparative investigation of the speech-associated attitude of preschool and kindergarten children who do and do not stutter. *Journal of Fluency Disorders*. 30 (4), 307–318.

Vanryckeghem, M., De Niels, T. & Vanrobaeys, S. (2015) The KiddyCAT: A test-retest reliability investigation. *Cross-Cultural Communication*. 11 (4), 10.

Verdon, S., Wilson, L., Smith-Tamaray, M. & McAllister, L. (2011) An investigation of equity of rural speech-language pathology services for children: A geographic perspective. *International Journal of Speech-Language Pathology*. 13 (3), 239–250.

Victorian Department of Human Services. (2004) *Children's Speech and Language Services: Eastern Metropolitan Region*. Melbourne, Victoria, Australia: Eastern Metropolitan Regional Office.

Vogel, A.P., Block, S., Kefalianos, E., Onslow, M., Eadie, P., Barth, B., Conway, L., Mundt, J.C. & Reilly, S. (2015) Feasibility of automated speech sample collection with stuttering children using interactive voice response (IVR) technology. *International Journal of Speech-Language Pathology*. 17 (2), 115–120.

Waite, M.C., Cahill, L.M., Theodoras, D.G., Busuttin, S. & Russell, T.G. (2006) A pilot study of online assessment of childhood speech disorders. *Journal of Telemedicine and Telecare*. 12 (3_suppl), 92–94.

Waite, M.C., Theodoros, D.G., Russell, T.G. & Cahill, L.M. (2010) Internet-based telehealth assessment of language using the CELF–4. *Language, Speech, and Hearing Services in Schools*. 41 (4), 445–458.

Waite, M.C., Theodoros, D.G., Russell, T.G. & Cahill, L.M. (2012) Assessing children's speech intelligibility and oral structures, and functions via an Internet-based telehealth system. *Journal of Telemedicine and Telecare*. 18 (4), 198–203.

Walker, J.P. (2015) *Speech Therapy Telepractice and Technology Program Manual.* University of Maine.

Wales, D., Skinner, L. & Hayman, M. (2017) The efficacy of telehealth-delivered speech and language intervention for primary school-age children: A systematic review. *International Journal of Telerehabilitation*, 9 (1), 55.

Ward, D. (1992) Outlining semi-intensive fluency therapy. *Journal of fluency disorders*. 17(4), 243-255.

Ward, D. (2018) *Stuttering and Cluttering: Frameworks for Understanding and Treatment*. Psychology Press.

Ward, E., Crombie, J., Trickey, M., Hill, A., Theodoros, D. & Russell, T. (2009) Assessment of communication and swallowing post-laryngectomy: A telerehabilitation trial. *Journal of Telemedicine and Telecare*. 15 (5), 232–237.

Ward, E., Sharma, S., Burns, C., Theodoros, D. & Russell, T. (2011) Using telerehabilitation to assess clinical dysphagia status. *Dysphagia*. 26 (4), 437–438.

Watkins, K.E., Smith, S.M., Davis, S. & Howell, P. (2007) Structural and functional abnormalities of the motor system in developmental stuttering. *Brain*. 131 (1), 50–59.

Watson, P.F. & Petrie, A. (2010) Method agreement analysis: A review of correct methodology. *Theriogenology*. 73 (9), 1167–1179.

Whitehead, E., Dorfman, V., Tremper, G., Kramer, A., Sigler, A. & Gosman, A. (2012) Telemedicine as a means of effective speech evaluation for cleft palate patients. *Annals of Plastic Surgery*. 68 (4), 415.

Wilkie, T. & Beilby, J. (1996) Post-treatment stuttering severity under different assessment conditions. *Australian Journal of Human Communication Disorders*. 24 (1), 19–27.

Williams, T.L., May, C.R. & Esmail, A. (2001) Limitations of patient satisfaction studies in telehealthcare: A systematic review of the literature. *Telemedicine Journal and e-Health*. 7 (4), 293–316.

Wilson, L., Onslow, M. & Lincoln, M. (2004) Telehealth adaptation of the Lidcombe program of early stuttering intervention: Five case studies. *American Journal of Speech-Language Pathology*. 13 (1), 81–92.

Wingate, M.E. (1964) A standard definition of stuttering. *Journal of Speech and Hearing Disorders*. 29 (4), 484–489.

Wittke-Thompson, J.K., Ambrose, N., Yairi, E., Roe, C., Cook, E.H., Ober, C. & Cox, N.J. (2007) Genetic studies of stuttering in a founder population. *Journal of Fluency Disorders*. 32 (1), 33–50.

World Health Organization. (1993) *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. World Health Organization.

World Health Organization. (1998) *Health21 – Health for All in the 21st Century*. Copenhagen, World Health Organization Regional Office for Europe.

World Health Organization. (2008) *Saudi Arabia Health Profile*. 3-8-2010. [Pamphlet].

World Health Organization. (2001) *ICF: International Classification of Functioning, Disability, and Health*. Geneva.

World Health Organization. (2015) *Process of translation and adaptation of ins*truments. Available from http://www.who.int/substance_abuse/research_tools/translation/en/ [Accessed 2nd March 2015].

Wilson, L., Lincoln, M. & Onslow, M. (2002) Availability, access, and quality of care: Inequities in rural speech pathology services for children and a model for redress. *Advances in Speech Language Pathology*. 4 (1), 9–22.

Wright, L. & Ayre, A. (2000). WASSP: A new outcome measure in adult stuttering therapy. *Journal of Fluency Disorders*. 25 (3), 222.

Wylie, K., McAllister, L., Davidson, B. & Marshall, J. (2018) Communication rehabilitation in sub-Saharan Africa: The role of speech and language therapists. *African Journal of Disability*. 7, 1–9.

Xue, S.A. & Lower, A. (2010) Acoustic fidelity of internet bandwidths for measures used in speech and voice disorders. *The Journal of the Acoustical Society of America*. 128 (3), 1366–1376.

Yairi, E. & Ambrose, N.G. (1992) A longitudinal study of stuttering in children: A preliminary report. *Journal of Speech, Language, and Hearing Research*. 35 (4), 55–760.

Yairi, E. & Ambrose, N.G. (1999) Early childhood stuttering I: Persistency and recovery rates. *Journal of Speech, Language, and Hearing Research*. 42 (5), 1097–1112.

Yairi, E. & Ambrose, N.G. (2005) *Early Childhood Stuttering for Clinicians by Clinicians*. Pro Ed.

Yairi, E. & Ambrose, N.G. (2013) Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*. 38 (2), 66–87.

Yairi, E., Ambrose, N.G. & Niermann, R. (1993) The early months of stuttering: A developmental study. *Journal of Speech, Language, and Hearing Research*. 36 (3), 521–528.

Yairi, E. & Seery, C.H. (2015) *Stuttering: Foundations and Clinical Applications*. Pearson.

Yaruss, J.S. (1997) Clinical measurement of stuttering behaviors. *Contemporary Issues in Communication Science and Disorders*. 24 (24), 33–44.

Yaruss, J.S. (2000) The role of performance in the demands and capacities model. *Journal of Fluency Disorders*. 25 (4), 347–358.

Yaruss, J.S. (2010) Assessing quality of life in stuttering treatment outcomes research. *Journal of Fluency Disorders*. 35 (3), 190–202.

Yaruss, J.S., Coleman, C. & Hammer, D. (2006) Treating preschool children who stutter: Description and preliminary evaluation of a family-focused treatment approach. *Language, Speech, and Hearing Services in Schools*. 37 (2), 118–136.

Yaruss, J.S., Coleman, C.E. & Quesal, R.W. (2012) Stuttering in school-age children: A comprehensive approach to treatment. *Language, Speech, and Hearing Services in Schools*. 43 (4), 536–548.

Yaruss, J.S. & Quesal, R.W. (2006) Overall assessment of the speaker's experience of stuttering (OASES): Documenting multiple outcomes in stuttering treatment. *Journal of Fluency Disorders*. 31 (2), 90–115.

Young, M.A. (1969) Observer agreement: Cumulative effects of rating many samples. *Journal of Speech, Language, and Hearing Research.* 12 (1), 135–143.

Zaki, R., Bulgiba, A., Ismail, R. & Ismail, N.A. (2012) Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review. *PloS One*. 7 (5), e37908.

Zebrowski, P.M. (2016) The high cost of stuttering: Insurance regulations impose huge barriers to stuttering treatment, even in university clinics. *The ASHA Leader*. 21 (7), 6–7.

Zebrowski, P.M. & Arenas, R.M. (2011) The “Iowa Way” revisited. *Journal of fluency disorders*, 36 (3), 144-157.

Appendices

Appendix 1 Questionnaire to Explore the Assessment Practice of Saudi Speech and Language Therapists

Section 1: Background Information

(1) Which category below includes your age?

o 21–29

o 30–39

o 40–49

o 50–59

o 60 or older

(2) What is your gender?

o Female

o Male

(3) In which city do you work?

…….....

(4) Where did you receive your degree?

……

(5) What is your highest level of academic achievement?

o Bachelor

o Masters

o PhD

(6) What is, currently, your most predominant work setting?

o In-patient acute hospital setting

o Out-patient acute hospital setting

o In-patient rehabilitation setting

o Outpatient rehabilitation setting

o Private practice

o Other: Please specify: …….

(7) In the course of your work, the approximate percentage of your caseload that contains clients who have stuttering is:

o 5%–10%

o 10%–25%

o 26%–50%

o 50%–75%

o 75%

(8) Please indicate your overall experience in speech and language therapy (months/years).

…….....

(9) Please indicate your experience with stuttering (months/years).

…….....

Section 2: Assessment Protocol for School-Age Children Who Stutter

(1) What professional is typically involved in the assessment of school-age children who stutter at your workplace?

……
……
……
……
……
……

(2) Are there any specified guidelines/protocols that you follow to assess school-age children who stutter?

o Yes

o No

If yes, please briefly describe

……

……

……

……

……

(3) What type of assessment do you use?

o Formal

o Non-formal

o Both

(4) In the case of formal assessment, please list the top three formal assessments that you most frequently use for school-age children who stutter:

1.

2.

3.

(5) In the case of formal assessment, how often do you adapt non-Arabic assessment tools?

o Never

o Rarely

o Occasionally

o Frequently

(6) Which adapted tools do you use most often?

……

……

……

(7) Which of these components do you usually include in your assessment of school-age children who stutter? (check all that apply):

o Frequency of stuttering (e.g. % syllables)

o Types of stutters

o Duration of stuttering

o Rate of speech

o Effect of task on stuttering (e.g. across different speech tasks such as reading aloud versus picture description)

o Effect of situation on stuttering (e.g. with friends versus with strangers)

o Speech naturalness (determining whether speech has a natural sounding quality)

o Secondary behaviours

o Covert stuttering behaviours (check all that apply):

* Emotional reactions

* Avoidance (avoidance of producing certain sounds or words or avoiding certain situation or people)

* Expectation of stuttering (the extent to which a stutterer believes that control and normal speech are impossible)

* Expectation of fluency (the belief that some form of control over speech is possible)

* Motivation

* Self-perception

(8) In addition to stuttering assessment, what other assessments do you carry out with school-age children who stutter? (Check all that apply):

o Oral motor function

o Language

o Articulation

o Phonology

o Other: please specify:

……

……

……

……

(9) Do you assess the impact of stuttering on the child's daily life and the attitudes of children who stutter?

o Yes

o No

If yes, how do you assess that?

……

……

……

……

……

(10) Is there any specific case history form you use for:

* School-age children who stutter?

o Yes

o No

If yes, please state the most important questions in your opinion

……

……

……

……

……

* Does it include questions for the parents?

o Yes

o No

If yes, please state the most important questions that address the parents in your opinion

……

……

……

……

……

(11) How long do you take to assess a school-age child who stutters?

……..

Section 3: Patient's Location

(1) How often do you see patients from outside the city you work in?

o Never

o Once or twice per year

o Once or twice per month

o Every week

(2) Do you think that patient performance would differ if they received speech and language therapy services in their own city/village?

o Yes

o No

Why?

……

……

……

……

(3) From your experience, how often do patients fail to attend their appointments due to travelling difficulties?

o Never

o Rarely

o Occasionally

o Frequently

Why?

……..……..……

……

……

Section 4: Telepractice

(1) Are you familiar with telepractice for speech and language therapy services?

o Yes

o No

(2) Have you used telepractice to provide speech and language therapy services?

o Yes

o No

(3) Would you consider using telepractice to provide speech and language therapy services?

o Yes

o No

Appendix 2 Procedure for Applying and Scoring the SSI-IV

The overall severity score is based on three components of the instrument that are derived from speech samples: 1) the percentage of syllables stuttered (%SS); 2) the duration of the longest three stutters within a sample; and 3) physical concomitants, which are based on physical behaviours related to the stutterer's speech. Conversion tables are provided with which to convert values for each component mentioned into 'task' scores. The three task scores are summed to give the total SSI score. These three measures are collected in different ways based on whether the child can read or not.

Speech samples are obtained from the children. For each child, a spontaneous speech sample and an age-appropriate reading text are obtained. If a child cannot read, the assessment is based on spontaneous speech samples only: one speech sample that is a conversation with the clinician and a second sample that is a conversation with one of the parents. All SSI-IV components are scored later from the recordings. During speech data collection with children who cannot read, a stimulus picture to elicit the sample is used (see Figure 19). Clinicians also try to ask, interrupt, and mildly disagree with the child to reflect normal conversations.

Figure 19 An Example of Picture Plates Used to Elicit Spontaneous Speech Samples

1- Frequency score

Frequency scores are of two types: frequency scores obtained from readers and frequency scores obtained from non-readers.

A- Frequency scores for non-readers:

The percentage of stuttered syllables (%SS) is obtained. Counts of all syllables in the sample and of stuttered syllables are obtained. Riley (1994) noted that syllable counting in real time is a difficult task; therefore, SLTs make a dot for each fluent syllable and a line for stuttered syllables and count them later. This procedure was followed in the current study.

The events that are counted as stuttered are defined as 'repetition or prolongation of sounds or syllables (including silent prolongations)' (Riley, 1994: p.4). In addition, there are behaviours that must be excluded from the definition of stuttering:

> 'Behaviors such as rephrasing, repetition of phrases or whole words, and pausing without tension are not counted as stuttering. Repetition of one syllable words may be stuttering if the word sounded abnormal (shortened, prolonged, staccato, tense, etc.); however, when these single-syllable words are repeated but are otherwise spoken normally, they do not qualify as stuttering using the definition just stated.' (Riley, 1994: p.4)

It is considered a single stuttering event when a child stutters on a syllable repeatedly. For example, 'I wa, wa, wa, wa, want to eat' has four syllables and one stuttered event. Stuttering events are expressed as a percentage of all syllables, which yields %SS. In cases involving multiple samples of the same speech type, the %SS is calculated for each speech sample, after which the average is obtained. The average is considered to be the raw %SS score. It is worth noting that Riley (1994) was precise about stutter and syllable count, which makes the test easy to administer. The frequency raw score is converted into a task score according to the table provided in the test. According to the manual, the participant is considered a non-reader when his/her reading ability is below third-grade level.

B- Frequency scores for readers:

The procedure followed for readers is similar to the procedure followed for non-readers, with one exception: readers have an additional reading task. Reading materials are designed to be appropriate for 8–9-year-olds, 10–11-year-olds, 12–13-year-olds, and adults. Each text is composed of approximately 200 syllables. Percentage of stuttered syllables is calculated as previously described. What distinguishes this procedure from the non-reader procedure is that

the frequency scores of the elicited and reading samples are separate. Reading and elicited samples are added to yield the overall frequency score.

2- Duration score

The duration score is the average of the three longest stutters within a sample, measured in seconds, for which a stopwatch is used for calculation. After averaging the three durations, the average is converted to a scale.

3- Physical concomitants

The physical concomitants score is based on observations of all the speaking samples that are scored from speech samples. Only observable phenomena associated with the stuttering should be scored. General behaviour such as anxiety is not rated. The categorisation of physical concomitants is as follows:

➢ Distracting sounds: non-speech sounds that accompany the stuttering (e.g. sniffing, blowing, and clicking sounds)

➢ Facial grimaces: any abnormal movement or tension about the face (e.g. pressing lips together tightly and blinking)

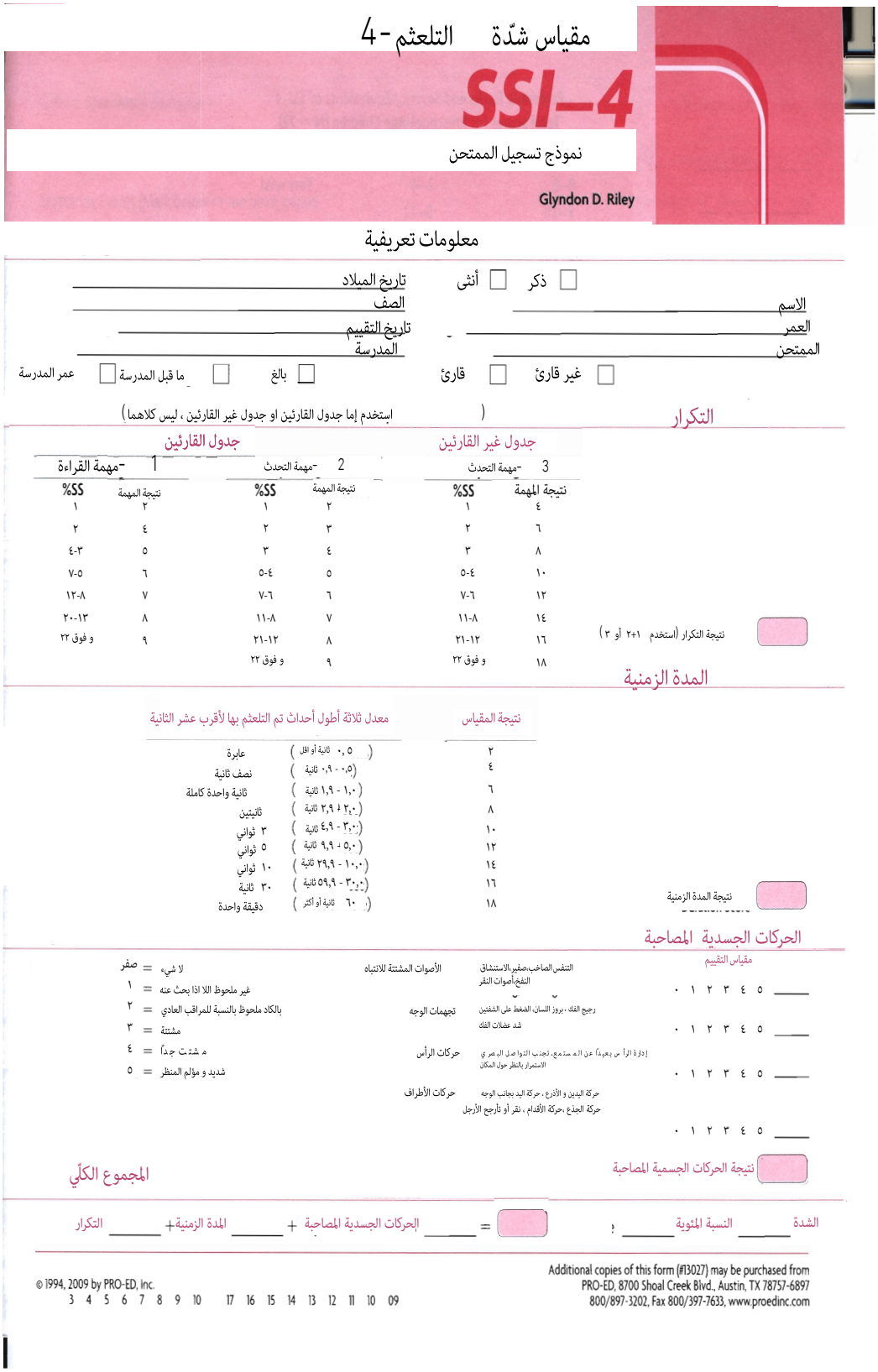
➢ Head movements: turning the head away from the listener to avoid eye contact, looking down at the feet, scanning the room, and looking up at the ceiling

➢ Movement of the extremities: any general body movement (e.g. foot tapping and arm swinging).

4- Total score

The total score is obtained by summing the scores of frequency, duration, and physical concomitants together. The relative stuttering severity can be ascertained by converting the total score to a percentile rank or severity equivalent.

Appendix 3 The Arabic Version of the SSI-IV



مقياس شدّة التلعثم-4

SSI–4

نموذج تسجيل الممتحن

Glyndon D. Riley

معلومات تعريفية

☐ ذكر ☐ أنثى تاريخ الميلاد

الاسم

الصف

العمر تاريخ التقييم

الممتحن المدرسة

☐ غير قارئ ☐ قارئ ☐ بالغ ☐ ما قبل المدرسة ☐ عمر المدرسة

التكرار

( ) استخدم إما جدول القارئين او جدول غير القارئين ، ليس كلاهما

جدول غير القارئين

3 -مهمة التحدث	
نتيجة المهمة	%SS
٤	١
٦	٢
٨	٣
١٠	٤-٥
١٢	٦-٧
١٤	٨-١١
١٦	١٢-٢١
١٨	و فوق ٢٢

جدول القارئين

2 -مهمة التحدث		1 -مهمة القراءة	
نتيجة المهمة	%SS	نتيجة المهمة	%SS
٢	١	٢	١
٣	٢	٤	٢
٤	٣	٥	٤-٣
٥	٤-٥	٦	٧-٥
٦	٦-٧	٧	١٢-٨
٧	٨-١١	٨	٢٠-١٣
٨	١٢-٢١	٩	و فوق ٢٢
٩	و فوق ٢٢		

نتيجة التكرار (استخدم ١+٢ أو ٣)

المدة الزمنية

نتيجة المقياس	معدل ثلاثة أطول أحداث تم التلعثم بها لأقرب عشر الثانية	
٢	(٠,٥ ثانية أو اقل)	عابرة
٤	(٠,٥ - ٠,٩ ثانية)	نصف ثانية
٦	(١,٠ - ١,٩ ثانية)	ثانية واحدة كاملة
٨	(٢,٠ + ٢,٩ ثانية)	ثانيتين
١٠	(٣,٠ - ٤,٩ ثانية)	٣ ثواني
١٢	(٥,٠ - ٩,٩ ثانية)	٥ ثواني
١٤	(١٠,٠ - ٢٩,٩ ثانية)	١٠ ثواني
١٦	(٣٠,٠ - ٥٩,٩ ثانية)	٣٠ ثانية
١٨	(٦٠,٠ ثانية أو أكثر)	دقيقة واحدة

نتيجة المدة الزمنية

الحركات الجسدية المصاحبة

مقياس التقييم

التنفس الصاخب،صفير،الاستنشاق النفخ،أصوات النقر ٠ ١ ٢ ٣ ٤ ٥ ____

تجهمات الوجه رجيج الفك ، بروز اللسان، الضغط على الشفتين شد عضلات الفك ٠ ١ ٢ ٣ ٤ ٥ ____

حركات الرأس إدارة الرأس بعيداً عن الهدف، تجنب التواصل البصري الاستمرار بالنظر حول المكان ٠ ١ ٢ ٣ ٤ ٥ ____

حركات الأطراف حركة اليدين و الذرع ، حركة اليد بجانب الوجه حركة الجذع ،حركة الأقدام ، نقر أو تأرجح الأرجل ٠ ١ ٢ ٣ ٤ ٥ ____

الأصوات المشتتة للانتباه

لا شيء = صفر

غير ملحوظ اللا اذا بحث عنه = ١

بالكاد ملحوظ بالنسبة للمراقب العادي = ٢

مشتتة = ٣

مشتت جداً = ٤

شديد و مؤلم المنظر = ٥

نتيجة الحركات الجسمية المصاحبة

المجموع الكلّي

الشدة _____ النسبة المئوية _____ ؛ _____ = _____ الحركات الجسدية المصاحبة + _____ المدة الزمنية + _____ التكرار

Appendix 4 Information Sheet

Information Letter for Participants

Research Project: Telepractice Application for the Clinical Assessment of Children Who Stutter

Supervisor:	Email:	Phone:
Dr David Ward	d.ward@reading.ac.uk	+44(0)1183784689
Experimenter:		
Lamya Aldukair		L.M.A@pgr.reading.ac.uk

My name is Lamya Aldukair, and I am completing my Doctor of Philosophy (PhD) at the University of Reading. I am carrying out a research project, and I am writing to ask you if you would be willing for your child to take part. The study aims to investigate the validity and reliability of using a telepractice application to assess stuttering in children. Before you decide whether you would like to take part, it is important that you understand why the research is being done and what it would involve for you. Please take the time to read the following information carefully and discuss it with others if you wish. Ask me if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

As professionals working in the field of speech and language therapy in Saudi Arabia, we are faced with many difficulties because there is a shortage of services for children who stutter, and there is a need to find alternative modes of service delivery in clinical settings to address the ever-increasing needs of the future. The present research aims, through a series of studies, to investigate the validity and reliability of using telepractice to assess stuttering in children as an alternative to face-to-face assessment. In addition, I am interested in knowing how satisfied the child and the parents are with assessment via telepractice.

Why have I been invited?

You have been invited to take part because you are the parent of a child who stutters. In my study, I will investigate how children who stutter perform on an assessment conducted from distance via telepractice, and I will compare the findings to assessments of other children who stutter who have been assessed face-to-face. In total, I aim to assess approximately 35 children, whose parents provide consent and who are happy themselves to participate.

Do I have to take part?

It is entirely your decision whether to take part in this research. If you decide that you are happy for your child to take part, you will be asked to sign a consent form. If you agree, you will still be free to withdraw at any stage of the project without the need to provide a reason.

What will happen if I take part?

If you decide that you are happy for your child to take part, then I will liaise with your child's clinician and arrange a convenient time to carry out the assessment during the day of any booked appointment with the clinician, thus ensuring that minimal disruption is caused.

The assessment will be completed in one session, of 30 minutes to one hour. At the beginning, I will tell your child what we will be doing and ask if they are happy to go ahead. If he/she agrees, I will assess your child's stuttering either from distance using video conferencing software or face-to-face.

I will video record each child's responses for further analysis.

What are the possible disadvantages and risks of taking part?

There are no disadvantages or risks in taking part in this assessment. However, if, at any point, a child indicates that they do not want to continue with the session for any reason, then the session will stop, with no consequences whatsoever.

What are the possible benefits of taking part?

Most children will find the assessment tasks fun and will be happy to participate, but the research will have no direct benefit for the individual children. The benefits will be for the wider community, as proving the proposed method's reliability will be beneficial for children who stutter by offering an alternative service model to improve access to services throughout the country.

What will happen when the research study stops?

After the project is completed, the participants' anonymous data and video recordings will be kept for five years and then destroyed, in line with university requirements.

Will taking part in the study be kept confidential?

All information in relation to this study will be kept confidential. A code number for your child will be used for all data (paper, computer, and video recordings), and therefore no personal information will be displayed. Additionally, video-recorded data will be kept in a locked cabinet and access to computer files will be available by password only. Recordings will be kept for five years, as required by the university.

What will happen to the result of the research study?

A report of the findings of the study can be provided to you if you are interested in the results. Anonymous results will be used by me and my supervisors, David Ward and Theo Marinis. No personal information will be available to anyone, besides me and my supervisors. If this research is published, anonymity will be maintained.

What do I have to do?

If you are happy to take part, please fill in and sign the consent form with this letter and return it either to your child's clinician or to me. If you would like further information, please do not hesitate to contact me or my supervisors (see contact details below).

I would be extremely grateful for your support, as taking part in this project will not only help me personally with my project but will also contribute to offering an alternative service delivery model to children who find it difficult to access speech and language therapy services, particularly children who stutter.

This application has been reviewed by the University Research Ethics Committee and has been given a favourable ethical opinion for conduct.

All investigators on this project have had criminal record checks and have been approved by the school to work with children.

Thank you for taking the time to read this information sheet.

Best wishes, Lamya Aldukair,

Appendix 5 Institutional Review Board (IRB) Approval

Kingdom of Saudi Arabia
King Saud University (034)
P.O. Box 7805 Riyadh 1172
Tel.: +966 11 4670011
Fax: +966 11 4671992
http://medicalcity.ksu.edu.sa

المملكة العربية السعودية
جامعة الملك سعود (٣٤)
ص.ب ٧٨٠٥ الرياض ١١٤٧٢
هاتف: ٠١١ ٤٦٧٠٠١١ +٩٦٦
فاكس: ٠١١ ٤٦٧١٩٩٢ +٩٦٦

جامعة الملك سعود
King Saud University

المدينة الطبية الجامعية

20.04.2016 (13.07.1437)
Ref. No. 16/0212/IRB

To: **Ms. Lamya Aldukair**
PhD Student
School of Psychology and Clinical Language Sciences
University of Reading – Whitenights
Reading RG6 6AL United Kingdom

Teacher Assistant at King Saud University
College of Applied Medical Science
Speech and Hearing Department
Email:

Cc: Dr. David Ward
Supervisor

Subject: **Research Project No. E-16-1867**

Project Title: "Telepractice Application for the Clinical Assessment of Children who stutter"

Dear Ms. Lamya Aldukair,

I am pleased to inform you that your above-mentioned research project was reviewed by the Institutional Review Board on 20 April 2016 (13 Rajab 1437). The project was **approved**. Work on this project may begin.

We wish you success in your research and request you to keep the IRB informed about the progress and final outcome of the study in a regular basis. Please quote the project number shown above in any future correspondence or follow-ups related to this study.

If you have any question, please feel free to contact me.

Thank you!

Sincerely yours,

لجنة أخلاقيات البحوث الطبية
College Medicine
Institutional Review Board (IRB)

Prof. Khalid M. Al-Faleh
MBBS, MSc, FRCPC, FAAP
Chairman, Institutional Review Board
King Saud University College of Medicine
King Saud University Medical City
P.O. Box 7805 Riyadh 11472 K.S.A.
E-mail:

الرقم
التاريخ /rubie
المرفقات

Appendix 6 Parents' Consent Form

Consent Form

Title of Study: Telepractice Application for the Clinical Assessment of Children Who Stutter

Please tick initial box

1.	I agree to take part in the above University of Reading research project. I have had the project explained to me, and I have read the participant information sheet, which I may keep for my records. I understand this will involve: Allowing my child to participate in an assessment of stuttering, at a time convenient to the clinic. Allowing my child's responses to be video recorded for later checking, and for the video recordings to be kept for five years as required by the University.	
2.	I understand that any information I provide is confidential, and that no information that could lead to the identification of any individual will be disclosed in any reports on the project or to any other party. No identifiable personal data will be published. The identifiable data will not be shared with any other organisation. I understand that anonymous data from this project (with no	

	names or personal details) will be shared with researchers at the University of Reading.	
3.	I understand that my participation is voluntary, that I can choose not to participate in part or all of the project, and that I can withdraw at any stage of the project without being penalised or disadvantaged in any way.	
4.	I agree to the University of Reading recording and processing this information about me. I understand that this information will be used only for the purpose(s) set out in this statement and my consent is conditional on the University complying with its duties and obligations under the Data Protection Act 1998.	
5.	I agree to take part in the above study.	

_____	_____	_____
Name of Participant	Signature	Date

Appendix 7 Child Consent Form

Child Consent Form

	Yes √	No X
Has somebody explained this project to you?		
Do you understand what this project is about?		
Have you asked all the questions you want to?		
Do you understand that it is okay to stop taking part at any time?		
Are you happy to be part of this study?		
Are you happy to do these activities with me?		

Your Name: _____

Today's Date: _____

My Name: _____

Appendix 8 Questionnaires

Parents' Opinion of Telepractice Pre-Assessment

As a means of improving the provision of services to patients in the future, you are requested to give your honest views of your child being assessed through an internet application. There are no right or wrong answers in this questionnaire, and your responses and identity will remain confidential. The questionnaire will take approximately ten minutes of your time to complete; there is space provided at the end for you to make comments pertaining to this research project. Once again, I thank you for your participation.

	Strongly Disagree	Disagree	Unsure	Agree	Strongly Agree
1. I will be happy to use telepractice if it is available in the hospital or healthcare facility nearest to my place of residence.	1	2	3	4	5
2. I am comfortable for my child to undergo stuttering assessment via the internet.	1	2	3	4	5
3. I will be happy for my child to be online and would consider using the internet for the stuttering treatment.	1	2	3	4	5
4. I believe my child will enjoy carrying out assessment tasks online.	1	2	3	4	5
5. I will have no difficulty in seeing the online speech therapist via the telepractice video link.	1	2	3	4	5

6. I will have no difficulty hearing the online speech therapist via the telepractice video link.	1	2	3	4	5
7. I would rate the online assessment as being equal to an assessment conducted traditionally in the face-to-face method.	1	2	3	4	5
8. The instructions given during the online assessment will be clear and easy to follow.	1	2	3	4	5
9. My child will have sufficient time to follow the instructions given during the assessment.	1	2	3	4	5
10. Telepractice can satisfactorily replace face-to-face assessment.	1	2	3	4	5
11. Telepractice will allow easy access to assessment and therapy for children who stutter.	1	2	3	4	5
12. Telepractice will save me travelling time and money.	1	2	3	4	5
13. I would prefer for my child to have a traditional (face-to-face) assessment with the speech therapist.	1	2	3	4	5

Please provide additional comments and/or suggestions that you may have with regard to this research project

_____

_____

Thank you for your participation

Parents' Opinion of Telepractice Post-Assessment

As a means of improving the provision of services to patients in the future, you are requested to give your honest views of your child being assessed through an internet application. There are no right or wrong answers in this questionnaire, and your responses and identity will remain confidential. The questionnaire will take approximately ten minutes of your time to complete; there is space provided at the end for you to make comments pertaining to this research project. Once again, I thank you for your participation.

	Strongly Disagree	Disagree	Unsure	Agree	Strongly Agree
1. I am happy to use telepractice if it is available in the hospital or healthcare facility nearest to my place of residence.	1	2	3	4	5
2. I was comfortable for my child to undergo stuttering assessment via the internet.	1	2	3	4	5
3. I was happy for my child to be online and would consider using the internet for the stuttering treatment.	1	2	3	4	5
4. I believe my child enjoyed carrying out assessment tasks online.	1	2	3	4	5
5. I had no difficulty in seeing the online speech therapist via the telepractice video link.	1	2	3	4	5
6. I had no difficulty hearing the online speech therapist via the telepractice video link.	1	2	3	4	5
7. I would rate the online assessment as being equal to an assessment conducted traditionally in the face-to-face method.	1	2	3	4	5

8. The instructions given during the online assessment were clear and easy to follow.	1	2	3	4	5
9. My child had sufficient time to follow the instructions given during the assessment.	1	2	3	4	5
10. Telepractice can satisfactorily replace face-to-face assessment.	1	2	3	4	5
11. Telepractice can allow easy access to assessment and therapy for children who stutter.	1	2	3	4	5
12. Telepractice will save me travelling time and money.	1	2	3	4	5
13. I would prefer for my child to have a traditional (face-to-face) assessment with the speech therapist.	1	2	3	4	5

Please provide additional comments and/or suggestions that you may have with regard to this research project.

_____

_____

Thank you for your participation

Patient's Opinion of Telepractice Pre-Assessment (School-Age)

As a means of improving the provision of services to patients in the future, you are requested to give your honest views of being assessed through an internet application. There are no right or wrong answers in this questionnaire, and your responses and identity will remain confidential. The questionnaire will take approximately ten minutes of your time to complete; there is space provided at the end for you to make comments pertaining to this research project. Once again, I thank you for your participation.

	Strongly Disagree	Disagree	Unsure	Agree	Strongly Agree
1. I will be happy to use telepractice if it is available in the hospital or healthcare facility nearest to my place of residence.	1	2	3	4	5
2. I am comfortable to undergo a stuttering assessment via the internet.	1	2	3	4	5
3. I anticipate I will be happy being online and would consider using the internet for stuttering treatment.	1	2	3	4	5
4. I anticipate I will enjoy carrying out assessment tasks online.	1	2	3	4	5
5. I anticipate I will have no difficulty in seeing the online speech therapist via the telepractice video link.	1	2	3	4	5
6. I will have no difficulty hearing the online speech therapist via the telepractice video link.	1	2	3	4	5

7. I would rate the online assessment as being equal to an assessment conducted traditionally in the face-to-face method.	1	2	3	4	5
8. The instructions given during the online assessment will be clear and easy to follow.	1	2	3	4	5
9. I expect to have sufficient time to follow the instructions given during the assessment.	1	2	3	4	5
10. Telepractice can provide a satisfactory alternative to face-to-face assessment.	1	2	3	4	5
11. Telepractice can allow easy access to assessment and therapy to children who stutter	1	2	3	4	5
12. I would prefer to have a traditional (face-to-face) assessment with the speech therapist.	1	2	3	4	5

Please provide additional comments and/or suggestions that you may have with regard to this research project.

_____

_____

Thank you for your participation

Patient's Opinion of Telepractice Post-Assessment (School-Age)

As a means of improving the provision of services to patients in the future, you are requested to give your honest views of being assessed through an internet application. There are no right or wrong answers in this questionnaire, and your responses and identity will remain confidential. The questionnaire will take approximately ten minutes of your time to complete; there is space provided at the end for you to make comments pertaining to this research project. Once again, I thank you for your participation.

	Strongly Disagree	Disagree	Unsure	Agree	Strongly Agree
1. I am happy to use telepractice if it is available in the hospital or healthcare facility nearest to my place of residence.	1	2	3	4	5
2. I was comfortable undergoing a stuttering assessment via the internet.	1	2	3	4	5
3. I was happy being online and would consider using the internet for stuttering treatment.	1	2	3	4	5
4. I enjoyed carrying out assessment tasks online.	1	2	3	4	5
5. I had no difficulty in seeing the online speech therapist via the telepractice video link.	1	2	3	4	5
6. I had no difficulty hearing the online speech therapist via the telepractice video link.	1	2	3	4	5
7. I would rate the online assessment as being equal to an assessment conducted traditionally in the face-to-face method.	1	2	3	4	5

8. The instructions given during the online assessment were clear and easy to follow.	1	2	3	4	5
9. I had sufficient time to follow the instructions given during the assessment.	1	2	3	4	5
10. Telepractice can satisfactorily replace face-to-face assessment.	1	2	3	4	5
11. Telepractice will allow easy access to assessment and therapy for children who stutter.	1	2	3	4	5
12. I would prefer to have a traditional (face-to-face) assessment with the speech therapist.	1	2	3	4	5

Please provide additional comments and/or suggestions that you may have with regard to this research project.

_____

_____

Thank you for your participation

Appendix 9 Types of ICC Models

ICC Type	Description
ICC (1,1)	Each subject is assessed by a different set of randomly selected raters, and reliability is calculated from a single measurement. Uncommonly used in clinical reliability studies.
ICC (1,k)	As above, but reliability is calculated by taking an average of the k raters' measurements.
ICC (2,1)	Each subject is measured by each rater, and raters are considered representative of a larger population of similar raters. Reliability is calculated from a single measurement.
ICC (2,k)	As above, but reliability is calculated by taking an average of the k raters' measurements.
ICC (3,1)	Each subject is assessed by each rater, but the raters are the only raters of interest. Reliability is calculated from a single measurement.
ICC (3,k)	As above, but reliability is calculated by taking an average of the k raters' measurements.

This table has been adopted from the University of Vermont website (Intraclass Correlation Coefficient, 2013)

Appendix 10 Summary of Audio and Visual Ratings of Sessions

Participant ID	Condition	Technical Difficulties Encountered	Overall Rating of Quality of Recording (poor=1; fair=2; good=3; very good=4; excellent=5)
1	TP-led*	Delay in signal reception, picture freezing, picture skipping, loss of image, and interruptions in audio	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
2	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
3	TP-led*	Delay in signal reception, picture freezing, picture skipping, loss of image, and interruptions in audio	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
4	Face-to-face-led		Audio: 1 2 3 4 5 Video: 1 2 3 4 5
5	TP-led*	Delay in signal reception, picture freezing, picture skipping, loss of image, and interruptions in audio	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
6	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
7	TP-led	None. However, the	Audio: 1 2 3 4 5

		participant's voice was extremely low	Video: 1 2 3 4 5
8	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
9	TP-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
10	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
11	TP-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
12	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
13	TP-led	Picture freezing	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
14	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
15	TP-led	Picture freezing, skipping, and pixilation	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
16	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
17	TP-led	Picture freezing, skipping, and pixilation	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
18	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
19	TP-led	Picture freezing	Audio: 1 2 3 4 5

			Video: 1 2 3 4 5
20	Face-to-face-led	Picture pixilation	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
21	TP-led	Picture freezing and pixilation	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
22	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
23	TP-led	Picture skipping and freezing	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
24	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
25	TP-led	Delay in sound transmission and picture pixilation	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
26	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
27	TP-led	Delay in sound transmission and picture pixilation	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
28	Face-to-face-led	None	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
29	TP-led*	Picture freezing, skipping, pixilation, sound transmission delay, and interruptions in audio	Audio: 1 2 3 4 5 Video: 1 2 3 4 5
30	Face-to-face-led		Audio: 1 2 3 4 5 Video: 1 2 3 4 5

*Note: The highlighted samples were severely compromised in terms of audio and video

Appendix 11 Additional Parents' Comments

PCH1: This model allows the service to reach people who can't travel and can use technology; also, I think this model allows for more privacy and makes people talk more since they will be communicating from a relaxed environment, from their homes. I also found the instructions given during the session to be very clear.

PCH2: I prefer the telepractice option due to transportation issues. I let my older son leave college early to drive me to the hospital. Also, I found the telepractice method flexible; it will save time, and the quality of the service is the same.

PCH3: I prefer face-to-face as the first session to establish rapport; at a later stage, I'll be happy to use the telepractice model.

PCH4: An advantage of this model is saving time; I don't need to be with my child. However, a main disadvantage is that the session might be affected by a weak internet connection.

PCH5: I am very pleased with this model; it is essential to follow technology advancements and incorporate such a service. The telepractice model can help services reach as any people as possible; my daughter engaged perfectly with the distant clinician.

PCH6: This model, in theory, is good, but our session was affected by slow internet.

PCH7: It is more convenient and has the same quality of service. Also, in this model, the child might be more relaxed. My son's clinician, Hakeem, travelled abroad to continue his studies; it would be brilliant if my child could continue therapy using this model with Hakeem; my son really progressed with him.

PCH8: Very interesting model; I am happy that we are following technological advances; I would definitely use it for my son's therapy.

PCH9: Convenient and ensures regular sessions, as his father is very busy, and I can't take him to therapy myself as I don't have a driver.

PCH10: I think this model is successful and promising as it uses technology to overcome distance barriers and offer services. In my case, I would happily use it for my son, especially since I live 50 km outside of Riyadh, where the nearest speech therapy clinics are.

PCH11: I have never seen my child that focused and engaged with the clinician in the face-to-face [sessions]; this model is flexible as it doesn't require my presence; also, it saves time and effort to travel, as, in my case, it requires four hours to go the appointment, attend it, and go

back home. Also, I think it is a positive thing that my son can utilise technology in something helpful like therapy.

PCH12: I think this model is convenient, as it saves time and effort; also, using this model will offer a more relaxed environment for my child because he sees the clinician in his natural environment [at home] and therefore the stress is less.

PCH13: Very positive experience; I would love to use it in the future, as it's very time saving and is similar to a regular service in the clinic.

PCH14: It is very good, and I would try it; I am also wondering if it's possible to hold group meetings with other stutterers using this model to discuss bullying?

PCH15: I think it's a great option to have the combination of face-to-face and telepractice. We can have telepractice session at times we can't make it to the clinic.

Appendix 12 Children's Additional Comments

CH1: I came from outside Riyadh to attend the appointment; this option will save me time. Although I do not have internet in my house, I can go to my grandparents' house next door or the closest hospital to my house.

CH2: I was less stressed because you were not with me in the same room

CH3: I preferred the telepractice method over the face-to-face because it felt less stressful seeing the therapist in a different location and through the screen. My hands usually sweat from stress when I talk to my SLT.

CH4: It is better than coming a long way to the hospital, especially for some people, and less stressful. Also, this model was better than what I expected; the sound and picture of the clinician were clear. Also, the pace of talking was slower, which made it a more relaxing experience [I think this was due to sound delays].

CH5: I liked the model even more when I tried it; it felt easy to communicate with the therapist and less stressful.

CH6: I felt less stressed, and it is easier to talk in this model.

CH7: It is very convenient and it saves travelling time; as I have the session within my home environment, I can avoid traffic.

CH8: This model was less stressful, and it was easier to communicate with the clinician. I can express myself more with this model.

CH9: I was surprised that I felt nervous talking to the clinician via the internet compared to a face-to-face session; I think it's because I get nervous too when I talk to someone over the phone.

CH10: I liked this model. However, it is new to me; I think I need to try it more to decide.

CH11: I'm neutral; I found the face-to-face and telepractice models to be the same; I don't mind going for both.

CH12: I liked it; it did not feel any different from a face-to-face session, except for moments of glitches, but it was OK.

CH13: Initially, before trying it, I thought they were the same; however, after the session, I think it was more fun because it incorporated technology; it is more convenient in terms of travelling to attend a session.

CH14: I loved it, and I would like to have more telepractice sessions in the future.