Assessing the oral language skills of bilingual preschool children with variable language experience: The exploration of a dynamic assessment approach

Wei Qin Teoh

BSpPath (Hons)

Speech Pathology

College of Nursing and Health Sciences

Flinders University

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

2 May 2019

Table of Contents

Abstract

Bilingual children living in multilingual countries are often exposed to a variety of languages in their home, school and community. Their language ability in each of the languages they speak is dependent on the amount of exposure and opportunities they have in each language. Therefore, bilingual children, even from the same bilingual community, often present with variable language experience. Consequently, the assessment of young bilingual children's oral language to screen for or diagnose language impairment becomes challenging.

Speech and language pathologists have to clinically determine whether a bilingual child with reported language concerns has a true language impairment, which requires intervention, or is a child who presents with perceived difficulties resulting from the lack of relevant language experience (i.e. language difference). Non-biased evaluations of bilingual children's language abilities are necessary to aid SLPs to differentiate between bilingual children with true LI and bilingual children with typical language development. Standardised assessments that are known as "tools of our profession" (Stow & Dodd, 2003. p. 363) may not provide an accurate assessment of bilingual children's language abilities as they are mostly designed and normed for monolingual children. Instead, a range of alternative assessment approaches such as dynamic assessment are recommended to be used for a less biased evaluation of bilingual children's language abilities. However, most of the recommendations stemmed from research that were conducted with bilingual children who formed the minority in predominantly monolingual English-speaking countries. It is unknown if speech and language pathologists working in predominantly bilingual countries are adopting recommended alternative assessment approaches when assessing bilingual children in their communities.

The aim of this thesis was to address the gap in the research on the use of alternative assessment approaches in evaluating the language skills of bilingual children from predominantly bilingual

countries. To do so, two phases of research were conducted in Singapore, an English-speaking but predominantly bilingual country in South-East-Asia. In Phase One of the research, a survey questionnaire was conducted to understand the current assessment approaches that are used by speech and language pathologists to assess the language skills of young Singaporean bilingual children; and the assessment challenges they faced. A total of 26 speech and language therapists working with the paediatric population in Singapore participated in the study. One of the key findings from Phase One was that the SLPs in Singapore were selecting commercially-available standardised assessments over recommended alternative assessment approaches when evaluating the language skills of Singaporean bilingual children.

To contribute to research on the use of alternative assessment approaches and as a continuation from Phase One, Phase Two of the study developed and explored the use of dynamic assessment process to assess the language skills of Singaporean bilingual children. Dynamic assessments assess children's language ability by evaluating their language learning potential in a 'Test-Teach-Retest' assessment format. The Validity Argument Framework (Kane, 2006; 2013) was adopted in Phase Two to gather a chain of favourable evidence from the development of the DA task to decisions based on the scores obtained on the DA. The performance on the DA of 48 typically developing English-Mandarin bilingual preschool children and 18 English-Mandarin bilingual preschool children who were identified to require on-going speech and language therapy were gathered. In addition, the performance of nine English-Mandarin bilingual preschool children with language concerns but were not identified to require speech and language therapist assessment was also gathered. These children were instead referred for learning support services with learning support educators. Evidence gathered found that in comparison to a standardised assessment and regardless of their language experience (i.e. English dominant or Mandarin dominant), the children's performance on the dynamic assessment was more accurate in differentiating those who were typically developing

and those who likely had language impairment (i.e. recommended to receive on-going speech and language therapy) than the standardised measure.

The findings from Phase One and Phase Two contribute to the understanding of current assessment approaches and challenges faced by speech and language pathologists working in a predominantly bilingual country. Most importantly, the findings provide validity evidence to advocate the use of dynamic assessment as a less biased assessment approach for assessing the language skills of bilingual children with variable language experience.

Declaration

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed: W. Q. Teoh

Date: 31 January 2018

Acknowledgements

Firstly, I would like to acknowledge my supervisors, Associate Professor Chris Brebner and Professor Sue McAllister. Thank you for all of your guidance, encouragement and wise words. The road towards a PhD is not just an academic one but a life changing journey. Besides the expertise you all have provided in the technical aspects of the research, you all have provided plenty of advice that enabled me grow into a stronger and better person throughout this journey.

Next, to my colleagues and management from the Department of Child Development, KKH, Singapore. Thank you for giving me this opportunity to pursue a postgraduate degree. The department and colleagues have provided a supportive platform for me to not only grow professionally as a clinician but also as an academic researcher.

I would like to also thank Enya, Amanda, Dr Sylvia Choo and Jennifer, the team at the DSP for supporting my research and assisting with the liaison with the preschools involved in the DSP. Thank you to the schools, principals, teachers and parents who consented to their children's participation in the study. I would also like to thank the Ministry of Social and Family Development, Singapore for granting permission for the study to take place in schools participating in the DSP.

Last but not least, a big thank you to my wonderful family and friends. To my husband, thank you for always believing in me and supporting me in each and every part of this journey. Thank you for being my rock and keeping everything in equilibrium while I fight my battles. To my parents and parents-in-law, thank you for always ensuring that I am well – always ensuring that I have plenty of rest and am eating well. To my siblings, thank you for always checking in with me. To my friends, thank you for the chats, the laughter, the meet-ups – they all kept me going on.

Dedication

To my angel in heaven – Bruce. To my son – Tobias.

Chapter 1 Introduction

1.1 Thesis Aims

The assessment of oral language skills in young bilingual[1] children to validly identify the presence or absence of language impairment (LI)[2] is challenging. LI is a form of development impairment where children without any known medical conditions or development disorders (i.e. Autism) demonstrate unexpected difficulties with the acquisition of the oral language (Tomblin et al., 1997; Kohnert, 2010). For bilingual children with LI, both languages are usually affected (Kohnert, 2010). Children with LI are consequently at risk of poor social, behavioural, educational and economic outcomes (i.e. Conti-Ramsden, Mok, Pickles & Durkin, 2013; Law, Rush, Schoon & Parsons, 2009; Parsons, Schoon, Rush & Law, 2011). Therefore, complete and valid assessment, followed by timely and responsive intervention is crucial for improving language and the related outcomes.

The main purpose of the study was to address the gap in the research on the use of alternative assessment approaches in evaluating the language skills of bilingual children from predominantly bilingual countries. This is to contribute to the existing literature on the use of appropriate assessment approaches in evaluating the language skills of young bilingual children with variable language experience and identifying those at risk of LI.

The research presented in this thesis was conducted in two phases (henceforth referred to as Phase One and Phase Two).

The primary aim of Phase One was to explore the current assessment approaches used by Speech and Language Pathologists (SLPs) in Singapore in evaluating the language skills of

[1] In this thesis, the term bilingual also refers to the term multilingual that is found in the literature. The term bilingual is defined as a person who has been exposed to two or more languages at home, school and in the community; and is able to understand and speak two or more languages in everyday conversations.

[2] In this thesis, the term LI also represents the following terms such as primary language impairment, specific language impairment and language disorder that are found in the literature.

Singaporean bilingual children. This was to have a better understanding of the challenges they may face in evaluating the language skills of Singaporean bilingual children and identifying those at risk of LI.

The primary aim of Phase Two was to develop and gather validity evidence on whether a dynamic assessment (DA) approach can validly differentiate between the language learning potential of children with and without LI, and therefore identify Singaporean bilingual preschool children at risk of LI. The Validity Argument Framework (Kane 2006, 2013) was adopted to guide the development of the DA process and to gather preliminary validity evidence. The intention was not to develop a finalised DA tool but to gather favourable evidence to explore, and if appropriate, support the proposed interpretations and uses of the DA process. The purpose of gathering validity evidence was to advocate the use of DA approach for evaluating the language skills of young bilingual children with variable language experience (Kane, 2006).

1.2 Research Significance

A growing culturally and linguistically diversified world has resulted in an increasing proportion of bilingual children being seen on speech and language therapy caseloads. The increased percentage of bilingual children seen are also apparent in predominantly monolingual countries (e.g. USA, UK, and Australia) where bilingual education is not compulsory (Hemsley, Holm & Dodd, 2014). At the same time, there are many countries in the world where bilingualism in the norm (e.g. Singapore, India, Canada). In these countries, bilingual education is compulsory, where children are mandated to learn English and an additional language based on their ethnicity or the region where they live. Therefore, the bulk of speech and language caseloads in these countries are made up of bilingual children speaking a diverse range of languages.

The valid assessment and identification of LI among young bilingual children can be challenging for two main reasons. Firstly, the rate of development in each language that a bilingual child speaks is dependent on the quality and quantity of input he or she has to each language (Hoff & Core, 2013; Kohnert, 2010). Therefore, the developmental trajectory in each language cannot be compared to monolingual developmental norms. At the same time, due to variability in input, it is highly likely that bilingual children who speak the same two languages from the same bilingual community will present with variable language abilities (Bedore & Peña, 2008; Kohnert, Windsor & Erbert, 2009; Paradis, 2010). Thus, it is often difficult to determine whether a bilingual child with language concerns has a true LI or language difference. Secondly, the research and development of language assessment tools, especially standardised assessments, has largely focused on monolingual children. Such assessments cannot be used to be validly screen or identify LI among bilingual children (De Lamo White & Jin, 2011).

In recent times, there has been an emerging body of research that focuses on understanding the challenges faced in assessing the language skills of young bilingual children (see Caesar & Kohler, 2007; Guiberson & Atkins, 2012; Williams & McLeod, 2012). At the same time, researchers have suggested practical alternative assessment approaches to be used (e.g. Bedore & Peña, 2008; De Lamo White & Jin, 2011; Ebert & Kohnert, 2016). Researchers have long advocated for the use of alternative assessment approaches such as DA and language sampling over commercially-available standardised assessments to evaluate the language skills of bilingual children (De Lamo White & Jin, 2011; Hemsley et al., 2014; Laing & Kamhi, 2003; Pieretti & Roseberry-McKibbin, 2016). There is also an emerging body of evidence demonstrating that such alternative assessment approaches can differentiate between bilingual children with and without LI with high classification accuracy (e.g. Gutiérrez-Clellen & Simon-Cereijido, 2009; Peña, Iglesias & Lidz, 2001; Peña, Gillam & Bedore, 2014). However, most of the research has been conducted in predominantly monolingual countries where

bilingual children make up a minority or small percentage of speech and language therapy caseloads.

Little is known about whether SLPs working in predominantly bilingual countries are adopting the suggested alternative assessment approaches to evaluate the language skills of bilingual children. It is also unknown if they face similar or different challenges in assessing bilingual children in their community. Similarly, there is also limited published research on the use of alternative assessment approaches to validly evaluate the language skills of bilingual children in predominantly bilingual countries.

1.3 Contextualising the Researcher's Position

The researcher's experience and challenges faced working as a bilingual SLP in her home country (Singapore), an English-speaking but predominantly bilingual country formed the basis of this study.

Firstly, I have observed that the lack of appropriate local language measures has resulted in SLPs relying on commercially-available standardised language assessment measures to assess the English language skills of Singaporean bilingual children. These commercially-available standardised language assessments are mostly designed and normed for use with monolingual English-speaking children in the UK, USA and Australia. It is often challenging to differentiate whether Singaporean bilingual children who performed below monolingual norms on commercially-available language assessments present with a perceived difficulty resulting from language difference or true language difficulty. Therefore, they are not appropriate to be used to assess the English language skills of Singaporean bilingual children. Local research has also yet to explore if SLPs in Singapore are using any other assessment approaches to validly assess the language skills of Singaporean bilingual children to assist in differential diagnosis.

It is also unknown if they face similar or different challenges in evaluating the language skills of bilingual children compared to their international counterparts.

Secondly, I was interested to explore if recommended alternative assessment approaches could validly evaluate the language skills of Singaporean bilingual children and therefore differentiate those who are typically developing from those at risk of LI. It was hoped that the research could be used as a stepping stone in contributing to the development of local assessment measures or techniques for use with Singaporean bilingual children.

1.4 Research Objectives

In Phase One, the assessment practices and challenges faced by SLPs working in a predominantly bilingual country, Singapore, were explored via a survey study.

The main research objectives were:

1. To explore whether SLPs working in Singapore were bilingual themselves.
2. To explore the current assessment approaches used by SLPs to evaluate the language skills of Singaporean bilingual children.
3. To explore the challenges faced by SLPs in assessing the language skills of Singaporean bilingual children.
4. To determine what assessment approaches SLPs will find useful in evaluating the language skills of Singaporean bilingual children.

One of the key findings from the survey study was that the SLPs in Singapore were selecting commercially-available standardised assessments over recommended alternative assessment approaches when evaluating the language skills of Singaporean bilingual children. The lesser to limited use of alternative assessment approaches could be attributed to the fact that no local study has explored the use of alternative assessment approaches with the Singaporean bilingual population. Therefore, due to the lack of research in this area, it was deduced that SLPs in

Singapore were likely not confident and/or comfortable in adopting alternative assessment approaches to evaluate the language skills of Singaporean bilingual children.

As a continuation from Phase One, DA was identified as an appropriate alternative assessment approach to be developed and explored in Phase Two of the study. DA assesses the children's language ability by evaluating their language learning potential in a 'Test-Teach-Retest' assessment format (Gutiérrez-Clellen & Peña, 2001). The purpose of Phase Two is to develop a DA process and explore whether it could validly evaluate the language learning potential of Singaporean bilingual children, and therefore identify those at risk of LI.

The Validity Argument Framework (Kane, 2006; 2013) was adopted to gather evidence for the proposed use and interpretation of the DA process. The purpose was to gather favourable evidence to support the intended interpretation and use of the DA: that when a child is observed to score poorly on the DA, the child is likely to have poor language learning potential and is at risk of LI. At this stage of development where the purpose is to explore the feasibility of DA process, the validity evidence to be gathered will be of confirmationist bias (Kane 2006; 2013; Chapelle 2008). Firstly, a chain of inferences, claims and assumptions from the development of the DA task to decisions based on the scores obtained on the DA was proposed (Kane, 1992; 2012). Next, the relevant evidence to support the claims under each inference was gathered.

Therefore, the research objective of Phase Two was to gather evidence for the proposed use and interpretation of the DA process using Kane's Validity Argument Framework (2006; 2013). Specifically, evidence was gathered to support the following claims:

1. Observations of performance on the DA reveal a bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI.

2. Observations of performance on the DA are evaluated to provide observed scores reflective of language learning potential.
3. Observed scores on the DA reflect language learning potential that are consistent among bilingual children with the same language developmental profiles regardless of their prior language experience.
4. The construct of language learning potential as assessed on the DA predicts scores that correspond to bilingual children's language developmental profiles.

1.5 Thesis Structure

This research is presented in 11 chapters, with this introduction comprising the first chapter.

Phase One of the study is presented in Chapter Two and Chapter Three. In Chapter Two, an overview of bilingual language development and the common challenges faced in identifying LI among bilingual children are first presented (Section 2.2 and Section 2.3). This is followed by a literature review of the common language assessment approaches used to evaluate the language skills of bilingual children (Section 2.4). Next, the research gaps in understanding the bilingual assessment practices and challenges faced by SLPs working in predominantly bilingual countries are elaborated upon (Section 2.5). This provides the basis for establishing the relevance and purpose for Phase One.

In Chapter Three, the methodology, results and findings from Phase One are presented. Firstly, the methodological process of data collection is outlined (Section 3.2). Specifically, the design of the survey questionnaire and the local context (i.e. Singapore) in which the study was conducted are elaborated upon. The results from the survey questionnaire that were completed by 26 SLPs working in Singapore are presented next (Section 3.3). Specifically, the assessment practices and challenges faced by SLPs working in Singapore are presented using descriptive statistics. Lastly, a detailed discussion of the current assessment approaches adopted by SLPs

in Singapore when assessing bilingual children and the assessment challenges they face in their practice are presented (Section 3.4). Last but not least, the study limitations from Phase One and the potential directions for future research are presented.

Phase Two of the study is presented from Chapter Four to Chapter Ten. In Chapter Four, a comprehensive literature review of DA and validation research is presented. The origin and concept of DA is first outlined (Section 4.2). This is followed by an elaboration on the two main DA approaches that are used for language testing (Mediated Learning Experiences: Feuerstein, Rand & Hoggman, 1979; Lidz, 1991; Graduated Prompting: Campione and Brown; 1987); the advantages and limitations in each approach are also elaborated upon (Section 4.3). Next, the research gaps in DA research are highlighted to provide the basis for understanding the research aims of Phase Two. A review of the literature on validation methods in language assessment development and the gaps on validation research for DA are presented in the subsequent sections (Section 4.5 and Section 4.6). The Validity Argument Framework by Kane (2006, 2013) was chosen to gather validity evidence from the development to the trialling of the DA process. The claims and assumptions under each inference (*Domain Definition, Scoring, Generalisation, Extrapolation*) of the Validity Argument Framework (Kane; 2006, 2013) for the proposed use and interpretation of the DA process are presented and elaborated upon in the final section of Chapter Four (Section 4.7).

The development of the DA process is presented from Chapter Five to Chapter Seven. Before the DA process could be developed, an appropriate language task had to be selected to be evaluated in the DA context. To select an appropriate language task, a thorough literature review was conducted. The process of reviewing the literature on the theoretical accounts of LI and the validity evidence for the eventual selection of a word learning task to be evaluated in the DA process (i.e. DA of word learning skills) are presented in Chapter Five. In Chapter Six, the key design components of the DA of word learning skills are elaborated upon. This is

followed by a description of a pilot study where the DA of word learning skills was trialled and revised to ensure that the scoring rubric of the DA of word learning skills could be objectively applied. The administrative procedure and scoring illustrations of the revised version of the DA of word learning skills that were used in Phase Two are presented in Chapter Seven.

The methodology to gather validity evidence to support the claims under each inference of the Validity Argument Framework for the proposed use and interpretation of the DA of word learning skills is presented in Chapter Eight. The specific aims and hypotheses are first presented in relation to the validity evidence that is required to be gathered to support the claims made under each inference of the Validity Argument Framework (Kane; 2006, 2013). This is followed by an elaboration of the selection of participants and the means of data collection.

Extensive analysis of the results is presented in Chapter Nine. The performance of typically developing Singaporean bilingual children and Singaporean bilingual children identified for on-going speech and language therapy on the DA of word learning skills and a commercially-available standardised assessment of vocabulary are first compared. In addition, the performance of a group of Singaporean children with language concerns on the DA of word learning skills are also presented.

In Chapter Ten, the results are discussed and presented as evidence to support the claims made under each inference of the Validity Argument Framework (Kane, 2006; 2013) (Section 10.2). In addition, when appropriate, the children's performance on the standardised assessment of vocabulary are compared to further substantiate the validity evidence for the DA of word learning skills. A discussion of how the research conducted in Phase Two can contribute to research on increasing the clinical acceptability of DA in bilingual clinical settings are presented next (Section 10.3). Finally, the limitations of Phase Two of the study and potential directions for future research are presented (Section 10.4).

For the final chapter of the thesis, Chapter Eleven, the key findings of Phase One and Phase Two of the study are summarised. Last not but least, the overall clinical implications from Phase One and Phase Two are presented.

Chapter 2 Phase One: Literature Review

2.1 Introduction

In this chapter, a review of the literature is presented. An overview of language development in bilingual children is first presented followed by an elaboration of the challenges faced in identifying LI among bilingual children. This provides the background for understanding the difficulties faced by SLPs in assessing the language skills of bilingual children. Next, the literature on common assessment approaches used by SLPs is presented. The advantages and possible drawbacks of each assessment approach is evaluated. Lastly, the research gaps regarding bilingual assessment practices are identified, providing the relevance and purpose for the conduct of Phase One of the study.

2.2 Language Development in Bilingual Children

Bilingual children learn and speak at least two languages and uses both languages in their day-to-day interactions. There are, however, differences in exposure patterns that influence their development in each language (Paradis, 2010). For example, there are bilingual children who are exposed to two or more languages at home and through early childcare experience from birth. This group of children are defined as simultaneous bilinguals. When provided with ample opportunities to learn and use each language, simultaneous bilinguals can become proficient speakers of both languages (Kohnert, 2010). There is another group of bilingual children who are exposed to only one language from birth and are exposed to the second language only after the age of 3;0 (years;months) when they commence formal schooling. This group of children are known as sequential bilinguals. Sequential bilinguals usually have their first language (L1) fairly well established before they begin to learn their second language (L2) in formal school settings (Bedore & Peña, 2008). Depending on the level of input in each language and the opportunities to use each language, sequential bilinguals may become proficient speakers in both languages eventually or in one language only (Kohnert, 2010).

For young and developing bilingual children, regardless of whether they are simultaneously or sequentially exposed to both languages, the levels of proficiency in both languages are determined by the amount of input and opportunities they have in their everyday learning environments (Bedore & Peña, 2008; Kohnert et al, 2009; Paradis, 2010). It is common for young bilingual children to be more dominant in one language over the other. The dominant language is usually the language that the child receives most exposure to. However, language dominance may shift over time if there are changes in the learning circumstances for each language (Kohnert, 2010). For example, a series of longitudinal case studies of Singaporean bilingual children by Gupta (1994) observed that many preschool children in Singapore were initially mostly exposed to their mother tongue in their home environment (i.e. Chinese children were expose to the Mandarin language). However when formal education commenced at primary school (i.e. six years old), there was shift in the use of home language to Standard Singapore English or Singapore Colloquial English to promote increase proficiency in English.

Besides considering the input and opportunities in each language, the development of both languages for bilingual children is often dependent on the sociolinguistic status of the two languages. When the two languages a bilingual child are exposed to are the official and majority languages of the country, he or she is likely to develop high levels of proficiency in both languages (Paradis, 2010). This is because both languages are commonly spoken and taught in the educational system, providing plenty of learning opportunities for the child to acquire the structures of both languages. On the other hand, when only one out of the two languages a bilingual child is exposed to is widely spoken and taught in the educational system, the bilingual child is likely to develop functional to high level of proficiency only in the majority language while the minority language is at risk for incomplete acquisition (Paradis, 2010).

The majority of bilingual children are typical learners. When provided with adequate input and opportunities to use both languages in their environments, bilingual children can become proficient speakers in both languages just like their typically developing monolingual counterparts in each language (i.e. Thordardottir, 2011; 2015). However, in the early stages of language development, the rate of acquisition in each language is highly variable. This is because individual and environmental circumstances often impact on the quality and quantity of input and opportunities to learn and use each language (Hoff & Core, 2013). Therefore, even for the same aged bilingual children who speak the same two languages, their rate of development in each language will be variable. Given the heterogeneous nature of young bilingual children's language development, it is often challenging for SLPs to determine whether a bilingual child with reported language concerns has a true LI, which requires intervention, or is a child who presents with perceived difficulties resulting from language difference, which does not. In the next section, the challenges in the identification of LI in bilingual children will be discussed.

2.3 The Challenges in Identifying Language Impairment among Bilingual Children

Globally, bilingual children now represent an increasing proportion of speech and language therapy caseloads (Hemsley et al., 2014). For bilingual children with LI, both languages are affected (Kohnert, 2010). Bilingual children with LI, like their monolingual counterparts with LI, experience difficulty in understanding and using languages and are at risk of poor academic, social and economic outcomes (i.e. Conti-Ramsden et al., 2013; Law et al., 2009; Parsons et al., 2011). Despite the increase in the number of developing bilingual children, there are still no epidemiological studies of LI in bilingual children (Kohnert, 2010).

The identification of LI in bilingual children is challenging for the following interrelated reasons. Firstly, the identification of LI in bilingual children is mostly based on language behavioural data in each language that is used to identify LI in monolingual children (Kohnert,

2010). Poor performance on language tasks is often considered as a behavioural indicator of LI. However, even within the monolingual population, children with LI form a mixed group with different profiles and severity of language difficulties (Conti-Ramsden & Durkin, 2011). These difficulties can be in the areas of vocabulary (Gray, 2004; Rice & Hoffman, 2015), morphosyntax (Conti-Ramsden & Windfuhr, 2002; Rice et al., 2010), narrative (Cleave, Girolametto, Chen & Johnson, 2010; Duinmeijer, De Jong, Scheper, 2012) and social language (Liiva & Cleave, 2005). The mixed nature of LI results in a lack of an agreed set of key inclusionary and exclusionary criteria for LI (Bishop, 2014; Reilly et al., 2014). The lack of definitive criteria for LI makes it even more challenging to identify LI in bilingual children as the rate of acquisition in each language even in same aged bilingual children who speak the same two languages is highly variable.

Secondly, there is a lack of local normative and language development data for the languages spoken in most bilingual communities and populations (Bedore & Peña, 2008). The fluid nature of language development among a heterogeneous group of bilingual children with changing language experience makes it almost an impossible task to study and record the developmental trajectories of the languages bilingual children speak as a group. Although bilingual children can acquire the structures of each of the languages they speak in the same sequence as their monolingual counterparts, the rate of language development in each language is influenced by the amount of language exposure they have in each language. For example, findings from several bilingual studies have found that the rates of vocabulary development (Hoff et al., 2012; Thordardottir, 2011; Pearson, Fernández & Oller, 1993) and morphosyntactic development (Brebner, McCormack & Rickard Liow, 2016; Paradis, 2010; Thordardottir, 2015) in bilingual children are influenced and determined by the amount of language input they have in each language.

Thirdly, SLPs often face difficulties in assessing both the languages spoken by bilingual children. In reality, SLPs often do not have the knowledge or proficiency of the diverse range of languages spoken within a bilingual community (Caesar & Kohler, 2007; Cruz-Ferreira & Ng, 2010; D'Souza, Bird, & Deacon, 2012; Williams & McLeod, 2012). As a solution, some researchers (D'Souza et al., 2012; Perrie & Core, 2006) have suggested increasing the number of bilingual SLPs to meet the demands of providing speech and language therapy services to bilingual children in both languages. However, it is uncertain if increasing the number of bilingual SLPs can resolve the challenges faced in assessing the language skills of bilingual children.

Lastly and most notably, in many bilingual communities, there is a lack of locally developed language assessment tools that can differentiate between typically developing and LI and therefore validly identify LI among bilingual children. Culturally and linguistically appropriate commercially-available assessment tools to discriminate between LI and difference among bilingual children are also rarely available (De Lamo White & Jin, 2011; Guiberson & Atkins, 2012; Stow & Dodd, 2003; Williams & McLeod, 2012). Assessment tools that assess the local languages spoken in the community are often not available.

Non-biased evaluations of bilingual children's language abilities are necessary to aid SLPs to differentiate between bilingual children with true LI and bilingual children with typical language development. There has been an emerging body of research on the common assessment approaches that SLPs can adopt when assessing bilingual children to evaluate the language skills of bilingual children and therefore assist with the identification of LI (e.g. Caesar & Kohler, 2007; De Lamo White & Jin, 2011; D'Souza et al., 2012; Laing & Kamhi, 2007; Williams &McLeod, 2012). The range of assessment approaches and strategies used by SLPs are discussed in the next section.

2.4 Assessment with Bilingual Children[3]

In this section, the use of common assessment approaches with bilingual children and their limitations will be elaborated upon. The purpose is to provide a summary of how these assessment approaches may or may not validly assess the language abilities of bilingual children and therefore assist in identifying LI. For a more detailed discussion of the critique of each assessment approach, readers may refer to De Lamo White & Jin's article (2011). An elaborative discussion of dynamic assessment is also provided in Chapter 4.

It is important to note that these assessment approaches alone may not yield a valid judgement of the bilingual child's language ability if they are not informed by other critical information. This includes the child's performance in all the languages he/she speaks while taking into consideration information about his/her language environment, patterns of language use and language dominance, first language abilities and developmental history gathered through case history interviewing, parent report and/or standardized questionnaires (Bedore et al., 2012; Brebner et al., 2016; De Lamo White & Jin, 2011; Paradis, 2016; Paradis, Emmerzael, & Duncan, 2010; Thordardottir, 2015).

2.4.1 Standardised Assessments

Surveys of the assessment approaches and practices used with bilingual children (e.g. Caesar & Kohler, 2007; D'Souza et al., 2012; Willams &McLeod, 2012) found that commercially-available standardised assessments are a common choice among SLPs when assessing bilingual children. Standardised assessments allow SLPs to evaluate a child's existing knowledge in one language and compare it to same age peers who are typically developing (De Lamo White & Jin, 2011). Most of these standardised assessments are created and standardised for specific

[3] This section was partially adapted from the original article that has since been published (Appendix 1). Wei Qin Teoh, Chris Brebner & Sue McAllister (2017): Bilingual assessment practices: challenges faced by speech-language pathologists working with a predominantly bilingual population, *Speech, Language and Hearing*, DOI: 10.1080/2050571X.2017.1309788

populations, usually monolingual speakers of the language. These assessments can only yield valid judgements if the bilingual child has similar language experience and exposure to the normative sample (De Lamo White & Jin, 2011). However, this is often not the case as language experiences among bilingual children are often variable. Furthermore, the stimulus materials and language targets of many standardised assessments do not match the cultural and linguistic experiences of bilingual children. Therefore, the resulting scores cannot be meaningfully interpreted or used to differentiate between typical developing or an impaired performance among bilingual children (Kohnert, 2010; Teoh, Brebner & McCormack., 2012).

Assessment adaptation/modification and re-norming have been explored to extend the use of commercially-available standardised assessments with bilingual populations, (e.g. Brebner, 2010; Okalidou, Syrika, Beckman & Edwards, 2011; Stokes & Wong, 1996). Although stimulus material and language output modifications may make standardised assessments more culturally and linguistically appropriate, the collection of normative data for the bilingual population is not only time-consuming but also impractical. Moreover, the collection of normative data can only be achieved with a relatively homogenous group whereas bilingual populations are more often heterogeneous in nature (De Lamo White & Jin, 2011; Ebert & Kohnert, 2017). As mentioned in previous sections (Section 2.2 and Section 2.3), the variability in language experience and dominance within each bilingual population leads to different rates and patterns of language acquisition and development even among same aged bilingual children who speak the same two languages (Brebner et al., 2015; Brebner et al, 2016; Teoh et al, 2017). Even if normative data are collected and available, it is impossible to determine if the bilingual child has the same language experience and exposure as the normative sample to be able to make a meaningful comparison.

Conceptual scoring (also known as composite scoring) has been employed as an alternative to single language scoring when assessing bilingual children's vocabulary or basic concept

knowledge on standardised assessments (Pearson, Fernandez & Oller, 1993; Gross, Buac & Kaushanskaya, 2014). Conceptual scoring counts the total number of vocabulary and basic concepts a bilingual child can express in any of the languages he or she is exposed to (Gross et al., 2014). Conceptual scoring is found to improve the scores of bilingual children to be comparable to that of monolingual children (Pearson et al., 1993; Thordardottir, Rothenberg, Rivard & Naves, 2006). It also reduces the over-identification of LI among simultaneous bilingual children (Gross et al., 2014). However, meaningful interpretation of conceptual scores is still problematic. Standardised assessments are designed to assess a single language hence translating test items into other languages for conceptual scoring purposes will not preserve the psychometric structure of the assessment (e.g. item difficulty hierarchy and the basal and ceiling rules of the standardised assessment) (Gross et al., 2016). Secondly, standardised assessments designed for specific populations can have culturally and linguistically biased items (De Lamo White & Jin, 2011). For example, in a standardised assessment designed for use in the USA, the picture of a 'totem pole' may be familiar to American children but may not be familiar to children from non-American cultures. Moreover, translated equivalents in other languages do not also always exist.

When commercially-available standardised assessments that are not designed or standardised for use with bilingual children are used, they should not be scored but instead be used to provide a description of a bilingual child's language ability (De Lamo White & Jin, 2011). However, this alone still does not provide sufficient information to allow SLPs to validly judge whether the bilingual child does or does not have a LI especially if only one language is evaluated. Besides the need to evaluate the second language, alternative assessment approaches should be used to supplement the description of the bilingual child's ability obtained from standardised assessments.

2.4.2 Criterion-referenced Measures

A criterion-referenced measure is an alternative assessment approach that evaluates a bilingual child's existing knowledge in a specific language domain (i.e. grammatical structures, linguistic concepts) in comparison with a predetermined performance criterion. The advantage of criterion-referenced measures is that SLPs can use stimuli, interaction patterns and contexts that are familiar to bilingual children hence minimising cultural and linguistic bias during evaluation (De Lamo White & Jin, 2011; Laing & Kamhi, 2003).

A bilingual child's performance is compared to the predetermined performance criterion based on a set of language performances that are expected of same aged bilingual children with similar language experience. This predetermined performance criterion is usually derived from local normative and language developmental data for that language spoken in the bilingual child's community (De Lamo White & Jin, 2011; Laing & Kamhi, 2003). However, in reality, there is often a lack of local normative and language development data for the languages spoken in most bilingual populations/communities. Therefore, it is often difficult for SLPs to set a valid predetermined criterion to determine whether the bilingual child is performing like his/her peers in a specific language domain or has a core language difficulty.

To overcome the lack or absence of well-established local normative and language developmental data to make a comparison to, some researchers have suggested that the predetermined performance criterion can be determined based on language patterns of parents or caregivers (Parent-Child Comparative Analysis: Terrell, Arensberg & Rosa, 1992). However, it has been argued that it is highly inaccurate to assume that linguistic behaviours of a child can be matched to an adult's (Laing & Kamhi, 2003). Besides the study conducted by Terrell and colleagues (1992), there is no other bilingual study published to date that has investigated the reliability and validity of using predetermined criterions that are based on the language patterns of parents or caregivers.

The research on the use of criterion-referenced measures to validly identify LI in bilingual children is still sparse. Therefore, SLPs seldom or rarely select the use of criterion-referenced measures to assess the language skill of bilingual children. This is evident from the results of the survey studies of bilingual assessment approaches and practices used with bilingual children (e.g. Caesar & Kohler, 2007; D'Souza et al., 2012; Williams & McLeod, 2012) where no single study reported the usage of criterion-referenced measures.

2.4.3 Language Sampling

Language sampling is one of the more common alternative assessment approaches adopted by SLPs for assessing the languages skills of bilingual children (D'Souza et al., 2012; Williams & McLeod, 2012). Language samples can be collected through narratives, picture description, and spontaneous conversations while using language, stimuli and interaction patterns that are familiar to the bilingual child (De Lamo White & Jin, 2011; Laing & Kamhi, 2003). Language samples can provide information on the child's language functioning, lexical knowledge and use of linguistic structures in both languages (e.g. Gutiérrez-Clellen & Simon-Cereijido, 2009; Ooi & Wong, 2012). For example, Ooi and Wong (2012) found that utterance length and syntax production information obtained from language sampling has the potential to provide diagnostic information to identify LI in Chinese-English bilingual Malaysian children.

Language samples should be collected in both/all the languages a bilingual child speaks (Gutiérrez-Clellen & Simon-Cereijido, 2009). SLPs can refer to established behavioural markers of LI or established language acquisition/development data for each language to assist in decision making. Bilingual children with LI should demonstrate difficulties in not one but both languages they speak (Bedore & Peña, 2008).

There are, however, challenges in using language sampling as an assessment approach to validly determine whether the child has LI. In reality, besides the English language, there is

still a lack of language developmental data and behavioural markers of LI for other languages for SLPs to make a comparison to (Bedore & Peña, 2008; Gutiérrez-Clellen & Simon-Cereijido, 2009; Ooi & Wong, 2012). Even if monolingual normative data for the language is available, some bilingual children, especially those with limited exposure to the language, may demonstrate behavioural markers of LI in that language that are secondary to language difference rather than a true language difficulty (Paradis, 2005; Thordardottir, 2015).

Furthermore, in the absence of an interpreter, language sampling analysis can also be challenging for SLPs who do not speak either or both languages spoken by the bilingual child.

2.4.4 Language-processing Measures

Language-processing measures are not a common alternative assessment approach used by SLPs despite being proposed as a least biased assessment (Williams & McLeod, 2012). Language processing measures evaluate children's underlying cognitive processing abilities that are used for processing and learning the language as well as other cognitive operations (Bedore & Peña, 2008; Paradis, 2010). Some examples of language processing measures are working memory tasks (e.g. digit recall, non-word repetition, word learning), perceptual tasks (e.g. discrimination of rapidly presented tones) and competing stimuli tasks (e.g. auditory figure ground, competing language processing task) (De Lamo White & Jin, 2011; Kohnert, Windsor & Yin, 2006; Laing & Kamhi, 2003).

Language-processing measures hold promise as a less biased assessment approach to validly identify LI in bilingual children. Unlike the other assessment approaches (e.g. standardised assessments, criterion-referenced measures, language sampling) which evaluate the children's existing language knowledge, children's performance on language processing tasks are not influenced or determined by their existing language knowledge (De Lamo White & Jin, 2011).

This is because language processing tasks evaluate children's processing functions that are used for learning the language.

However, findings (De Lamo White & Jin, 2011; Kohnert et al., 2006) have been mixed as to whether poor performances in language processing tasks such as sentence repetition and non-word repetition can validly be interpreted as associated with, or predictive of, LI in bilingual children (See Chapter Five, Section 5.2.2 for further discussion). There is, however, some emerging evidence that performance in word learning tasks can discriminate between bilingual children with LI and typically developing bilingual children (Kapantzolgou, Restrepo & Thompson, 2012). Nonetheless, researchers in the field agree that the diagnostic value of language processing measures to identify LI in bilingual children has not been fully explored and more research in this area is required (De Lamo White & Jin, 2011; Laing & Kamhi, 2003; Roseberry-McKibbin & O'Hanlon, 2005).

2.4.5 Dynamic Assessments

Dynamic assessments (DA) are becoming a preferred alternative assessment approach for assessing bilingual children (D'Souza et al., 2012; Williams & McLeod, 2012). DA evaluates the children's capacity to learn language targets in response to learning opportunities provided within the assessment itself. It allows SLPs to observe how children learn the language. It is also known to be a less biased assessment approach as unlike other knowledge-based assessment approaches (e.g. standardised assessments, criterion-referenced measures, language sampling), children's performances on DA are not affected by the environmental (i.e. low social economic status, lack of schooling) and linguistic variables (i.e. lack of exposure to the language, other varieties of the language).

DA draws on Vygotsky's model of cognitive development which proposes that learning takes place in the "Zone of Proximal Development (ZPD)" (Vygotsky, 1978; cited in Gutiérrez-

Clellen & Peña, 2001: 212). For the purposes of language assessment, this is defined as a language target that is developmentally achievable for the child. The goal of DA is to identify if the child is able to demonstrate the ability to learn language targets that are in their ZPD. In other words, the child's language learning potential.

There are several DA approaches (Refer to Chapter Four, Section 4.3). To discriminate language difference from LI, the 'Test-Teach-Retest' paradigm is often used. In the 'Test-Teach-Retest' paradigm, the SLP identifies a language area in which the child has difficulties and provides intervention in the form of Mediated Learning Experiences (Feuerstein, Rand & Hoffman, 1979; Lidz, 1991) or Graduated Prompting (Campione and Brown; 1987) to improve the child's functioning in the targeted area. The SLP then evaluates the child's performance again in the targeted area to identify changes that would indicate the child's ability to acquire language targets after teaching (Gutiérrez-Clellen & Peña, 2001). Children who show significant improvements are likely to have language difference due to the lack of familiarity or experience with the content assessed, whereas children who demonstrate limited changes are likely at risk of LI (Gutiérrez-Clellen & Peña, 2001).

DA has been shown to be useful in discriminating between typically developing bilingual children from bilingual children with LI in vocabulary and word learning studies (Camilleri & Botting, 2013; Kapantzoglou et al., 2012; Peña et al., 2001), narrative studies (Peña et al., 2006) and sentence structure studies (Hasson, Dodd & Botting, 2012). An advantage of DA is that the assessment can be carried out in a common language (e.g. English) spoken by both the child and SLP. This is especially useful when the SLP does not speak the dominant language of the child. Additionally, local normative and language developmental data for the languages spoken in the bilingual community are usually not required to make a comparison.

2.5 Research Gap

The assessment approaches summarised above were derived from the research and survey studies conducted in predominantly monolingual English-speaking countries (e.g. Australia: Williams & McLeod, 2012; Canada: D'Souza et al., 2012; USA: Caesar & Kohler, 2007; Laing & Kamhi, 2003; UK: Stow & Dodd, 2003). The incidence of bilingual children seen on SLPs' caseload in predominantly monolingual English-speaking countries has been increasing where the majority of the SLPs surveyed have reported bilingual children making up an increasing proportion of their caseloads (i.e. Caesar & Kohler, 2007; Williams & McLeod, 2012). Although the frequency of bilingual children seen on SLPs' caseloads has been increasing, the total number of bilingual children seen still only make up a small percentage of their caseloads. The majority of these SLPs' caseload is still made up of monolingual English-speaking children. Moreover, the majority of the SLPs in these countries were reported to be mostly monolingual speakers of the English language. For example, the surveys studies conducted in the USA and Australia found that only 6.2% (Caesar & Kohler, 2007) to 9.4% (Williams & McLeod, 2012) of these SLPs were sufficiently proficient in another language other than English. Some researchers (D'Souza et al., 2012; Perrie & Core, 2006) have therefore suggested increasing the number of bilingual SLPs so as to the meet the demands of providing speech and language therapy services to bilingual children in both languages.

On the other hand, there has been limited published research on how SLPs working in predominantly bilingual countries are evaluating the language skills of bilingual children in their communities. In these countries, the majority of the SLPs' caseloads are made up of bilingual children. Therefore, to contribute to the body of research on recommended assessment approaches to assess the language skills of bilingual children and therefore identify bilingual children at risk of LI, it is necessary to investigate the types of assessment approaches commonly used by SLPs working in a predominantly bilingual country. It is unknown whether

the majority of the SLPs working in predominantly bilingual countries are bilingual themselves, and if they are more likely to use recommended alternative assessment approaches or other assessment approaches when assessing bilingual children. Moreover, the type of assessment challenges SLPs faced in predominantly bilingual communities/countries are also not well understood.

2.6 Chapter Summary

An overview of the literature on bilingualism development, challenges in identifying LI among bilingual children and common assessment approaches used by SLPs to identify LI among bilingual children was undertaken. In the early stages of language development, the rate of development in each language among bilingual children is determined by the amount of exposure and opportunities they have to each language in their everyday context. Therefore, even the language abilities among bilingual children from the same community can be highly variable. The identification of LI among bilingual children is made even more complex as the behavioural indicators for LI are mostly based on language behavioural data available for monolingual speaking children in each language. In addition, the fluid nature of language development among bilingual children makes it challenging to study and record the developmental trajectories of the languages bilingual children speak and to identify potential behavioural indicators of LI.

To determine whether a bilingual child has LI, SLPs use a variety of assessment approaches to assess the language skills of bilingual children. Commercially-available standardised assessments that are designed and normed for use with monolingual children are not recommended to be used to assess the language skills of bilingual children. Instead, there is an emerging body of research that supports the use of alternative assessment approaches to assess the language skills of bilingual children.

The literature reviewed was based on bilingual research and survey on bilingual assessment practices conducted in predominantly monolingual countries. The research on bilingual assessment practices and challenges faced in identifying LI among bilingual children in predominantly bilingual countries have not been thoroughly explored. Therefore, Phase One of the study aims to investigate the assessment approaches used and challenges faced by SLPs working in a predominantly bilingual country. This is to contribute to the emerging body of research on the use of appropriate assessment approaches to validly assess the language skill of bilingual children.

In the next chapter, the methodology used and findings from Phase One of the study are presented.

Chapter 3 Phase One: Methodology to Discussion

3.1 Introduction

Phase One of the study aimed to investigate the assessment approaches used and challenges faced by SLPs working in a predominantly bilingual country. In this chapter, the methodology to the discussion of the findings from Phase One are presented. In the first section (Section 3.2), the context (Singapore) in which the study was conducted will be first described. Specifically, the bilingual context, incidence of LI and SLP services available in Singapore will be elaborated upon to provide an overview of the complexity involved in assessing the language skills of Singaporean bilingual children. Next, the specific research questions to be explored in Phase One are presented. Lastly, the means (i.e. survey questionnaire) in which data were collected and the recruitment of participants will be described.

In the second section of this chapter (Section 3.3), the analysis of the responses gathered from the survey are presented using descriptive statistics. The demographic information of the SLPs who participated in the survey compared to the demographics of children on their caseloads are first presented. This is followed by the reported analyses of the assessment approaches used, and challenges faced when assessing the language skills of Singaporean bilingual children. Finally, the SLPs' perceptions of the type of assessment approaches that can be locally developed and the assessment factors to consider during development are reported.

In the final section of this chapter (Section 3.4), the key findings of the study are presented. Firstly, the SLPs' and their caseload language match is discussed. Specifically, whether SLPs were bilingual themselves and if the languages they spoke matched with their caseloads' language demographics. Next, the assessment approaches used, and challenges faced when assessing the language skills of Singaporean bilingual children are discussed. The possible reasons for why certain assessment approaches are preferred over others are discussed and

presented. Lastly, based on the SLPs' perceptions on useful assessment approaches and factors to consider when developing local assessments, the researcher proposes potential development in the area. Finally, the limitations of the study and directions for future directions for research are discussed.

3.2 Methodology

An online survey instrument methodology was adopted to maximise the rate of participation. The online survey instrument allowed the researcher to invite all SLPs working in Singapore via an e-mail bulletin to participate in the study. SLPs who received the invitation and were keen to participate were then be able to complete the survey in their own time. A focus group methodology was considered but not adopted as it required SLPs to take time off their busy clinical schedules and it was anticipated that this would have limited the number of SLPs participating in the study. Likewise, other studies (Caesar & Kohler, 2007; D'Souza et al., 2012; Williams & McLeod, 2012) of similar nature have adopted survey instrument methodology to explore SLPs' assessment practices and challenges faced in assessing bilingual children in their community.

3.2.1 Ethical Clearance

Ethical clearance for the study to take place was obtained from the Social and Behavioural Research Ethics Committee, Flinders University of South Australia (Appendix 2).

3.2.2 The Singapore Context

3.2.2.1 Bilingual Context

The multilingual landscape in Singapore and its predominantly bilingual population makes it an appropriate site to investigate the bilingual assessment practices and challenges faced by SLPs working with a predominantly bilingual population. Singapore is a multicultural and multilingual country in Southeast Asia with an estimated population of 5.5 million people

(Singapore Statistics Board, 2014). At least 71% of the population (aged 15 and above) are literate in two or more languages (Singapore Statistics Board, 2010). Four official languages (English, Mandarin, Malay, and Tamil) and a variety of other non-official languages are spoken (Gupta, 1994). Two varieties of English are spoken in Singapore: Singapore Standard English and Singapore Colloquial English. Singapore Standard English is grammatically similar to Standard English spoken elsewhere but with a variation in accent. Singapore Colloquial English is described as having simple grammar and morphology while being pragmatically rich (Gupta, 1994). Language exposure within the population is highly variable as children can be simultaneously or sequentially exposed to two or more languages.

Bilingual education is compulsory in Singapore with English being the official language for education. Singaporean children are required to learn at least two languages, English (Singapore Standard English) and their mother tongue based on their ethnicity when formal education commences. English (Singapore Standard English) is the main and official language of instruction in schools where all subjects (i.e. Mathematics, Science, Arts and Crafts, Physical Education) are taught in. Students are grouped into their mother tongue classes based on their ethnicity (i.e. Chinese children are grouped together for their mother tongue class conducted in Mandarin; Malay children are grouped together for their mother tongue class conducted in Malay). The estimated percentage of English instructions versus mother tongue instructions used in school is 80% versus 20%. Despite English being used more frequently in school, close to 70% of Singaporeans still prefer to speak in their mother tongue at home and in the community (Department of Statistics Singapore, 2010). Even in homes where English is the preferred language, Singapore Colloquial English is used more widely as compared to Singapore Standard English, especially among everyday conversations with young children (Chua, 2011 & Gupta, 1994). High migration rates (Yeoh & Lin, 2012) have also resulted in

the introduction of other languages (i.e. Tagalog, Hindi) spoken within the community and this further diversifies the multilingual landscape in Singapore.

As a result of the complex language environment and mandatory bilingual education in Singapore, each Singaporean child is exposed to a unique language learning environment depending on the number, frequency and quality of languages used in their home, school and in the community. Therefore it is critical to consider each child's language exposure and dominant language as variability in language exposures will lead to different language profiles (Brebner et al, 2016). Local studies (Brebner et al., 2016; Teoh et al., 2012) conducted in Singapore have found that verb tense markings and vocabulary acquisition in the English language is impacted by language dominance Therefore, for the purpose of research, local studies (e.g. Brebner et al., 2016; Gn, Brebner, McCormack, 2014; Pua, Lee & Rickard-Liow, 2017; Teoh et al., 2012) conducted in Singapore have often grouped Singaporean bilingual children as English dominant (ED) or Non-English dominant (NED) based on parents' and teachers' questionnaire reports on the children's language exposure and use.

3.2.2.2 Incidence of Language Impairment in Singapore

There is no official figure on the incidence of LI in Singapore. LI is reported to affect both monolingual and bilingual children in similar numbers (Kohnert, 2010). The prevalence of LI among monolingual children in the USA is reported to be about 7.0% (Tomblin et al., 1997). Therefore, it is assumed that 7.0% of Singaporean bilingual children will have LI as well. In addition, LI is reported to be the single most common diagnosis seen in the Child Developmental Units in Singapore with 26.0% to 29.0% of new referrals (Age 7 and below) being diagnosed with isolated speech and language delay annually (Ho, 2007; Lian et al., 2012).

3.2.2.3 Speech and Language Therapy services in Singapore

The provision of speech and language therapy services in Singapore emerged formally in the 1980s and has since then expanded rapidly. The number of SLPs grew from a mere five therapists in the mid-1980s (Gupta & Chandler-Yeo, 1994) to a total number of 358 in 2014 (Allied Health Profession Council of Singapore, 2014). The demographics of SLPs practising in Singapore has also changed drastically. Currently, 72.0% of registered SLPs in Singapore (Allied Health Profession Council of Singapore, 2014) are Singaporean Citizens or Singaporean Permanent Residents while in the 1980s, the majority of SLPs working in Singapore were non-citizens (Gupta & Chandler-Yeo, 1994). The increase in Singaporean SLPs can be partially attributed to the introduction of a local graduate program in speech and language pathology in 2008 and the provision of government scholarships for Singaporean residents to access speech and language pathology programmes overseas. With the increase in qualified SLPs in Singapore, the number and types of speech and language therapy services available have also expanded rapidly.

Paediatric speech and language therapy services (0- to 12-year-old children) are provided in Child Development Units (0- to 7-year-old children) or/and Rehabilitation Units (0- to 12-year-old children) in three out of the eight restructured (i.e. public) hospitals in Singapore. Paediatric speech and language therapy services are also available in the community through government-run Voluntary Welfare Organizations (VWOs) (e.g., early intervention schools, therapy hubs that provide services) and privately-run clinics and/or early intervention centres.

A government-initiated program, the 'Development Support Program' (DSP) (Ministry of Social and Family Development Singapore, 2013), was introduced in 2013 where selected kindergartens and childcare centres can access onsite speech therapy services for pre-school children with mild developmental needs (7-year-old and below children). In the educational sector for school-aged children (7- to 12-year-old children), speech and language therapy

services are only available in international schools and special-needs schools (government-funded and private). There are no on-site speech and language therapy services provided in mainstream schools for school-aged children, but they can access services from therapy hubs in the community, rehabilitation units (i.e. government-funded hospitals) or privately run clinics/hospitals.

Although there has been a rapid increase in the number of SLPs and the provision of paediatric speech and language therapy services in Singapore, the assessment approaches used and challenges faced in assessing Singaporean bilingual children have not been explored and well understood. Therefore with such growth in the profession, it is timely to explore the current assessment approaches used by SLPs in evaluating the language skills of Singaporean bilingual and understand the type of assessment challenges they may face in their practice.

3.2.2.4 Research Questions

To explore the current assessment approaches used by SLPs in evaluating the language skills of Singaporean bilingual children and to understand the types of assessment challenges SLPs may face, the following research questions were asked:

1. Are SLPs working in Singapore themselves bilingual?
2. What are the assessment approaches currently used by SLPs to evaluate the language skills of Singaporean bilingual children? Are they more likely to use alternative assessment approaches compared to their international counterparts?
3. What are the assessment challenges faced by SLPs working in Singapore?
4. What assessment approaches, if developed for use with Singaporean bilingual children, will SLPs find useful in assisting them in evaluating the language skills of Singaporean bilingual children and identifying those at risk of LI?

3.2.3 Survey Design

A 16-item survey questionnaire was developed to explore the assessment approaches used and challenges faced by SLPs when assessing the language skills of Singaporean bilingual (Appendix 3). During the development of the survey questionnaire, the researcher reviewed survey questionnaires conducted internationally by other researchers on the topic of bilingual assessment approaches and practices (Caesar & Kohler, 2007; D'Souza et al., 2012; Guiberson & Atkins, 2012; Phoon & Maclagan, 2009; Williams & McLeod, 2012).

To ensure that the items in the survey questionnaire were relevant to local assessment practices and challenges; and were presented in a clear and concise manner, the survey questionnaire was piloted with five practising SLPs working in Singapore with varying years of practice (i.e. 2 to 8 years), nationalities (i.e. Singaporean, non-Singaporean) and practice settings (i.e. restructured hospital, VWOs, private). These five practising SLPs were sourced from the personal contacts of the researcher. The five SLPs completed the survey questionnaire and provided their feedback via email regarding the phrasing of the questions and appropriateness of the survey items. The feedback resulted in one minor content modification (i.e. creation of two separate questions on assessment practices with ED children and NED children due to variability in language experience among Singaporean bilingual children) and two minor presentation modifications (i.e. bolding of keywords in questions and reducing the number of questions on a single page).

The final survey questionnaire (See Appendix 3) comprised of questions about demographic information of SLPs (#3, #12 - #16), caseload demographics (#1 - #2), current assessment practices (#5 – #8), assessment challenges (#9) and SLPs' perception on the type of assessment approaches they deemed useful to be locally developed (#10 - #11). All questions had the option of free text boxes where SLPs could also input their opinions and/or comments.

3.2.3.1 SLPs' Demographic Information

SLPs were asked to specify their nationality (i.e. Singaporean or non-Singaporean), years of experience in the field, current and previous (if any) settings of practice, and countries in which they obtained their SLP qualifications. They were also asked to report the languages they spoke and to rate their proficiency in each language they spoke on a Likert-scale (i.e. Not at all, Minimal, Functional, Proficient).

3.2.3.2 Caseload Demographics

SLPs were asked to indicate the percentage of children on their caseloads for each age range (i.e. 0- to 6-year-old, 6- to 12-year-old) and percentage of children dominant in each of the official and majority languages (i.e. English, Mandarin, Malay, Tamil) spoken in Singapore.

3.2.3.3 Assessment Approaches

Due to the diversified language landscape in Singapore and for the purpose of research, local studies (i.e. Brebner et al., 2016; Gn et al., 2014; Teoh et al., 2012) conducted in Singapore have often grouped Singaporean bilingual children as ED or NED. Therefore similarly, this study sought to explore if there was a difference in the types of assessment approaches used when assessing ED or NED children.

SLPs were asked to rate the frequency with which they used each assessment approach (i.e. standardised assessments, language sampling, criterion-referenced measures, DA, processing-dependent measures) when assessing the language skills of Singaporean children who are ED and NED on a Likert-scale (i.e. Not at all, Rarely, Sometimes, Often, All the time). They were also asked to rate the frequency with which they used each commercially-available standardised assessment and the methods (i.e. compare a child's performance to normative data, estimate child's performance based on clinical experience etc.) they used to interpret the children's performances on standardised assessments. Lastly, they were also asked to rate the

frequency with which they used locally-developed language assessment measures on a Likert-scale.

3.2.3.4 Assessment Challenges

SLPs were asked to rate the frequency with which they encountered potential challenges when assessing the language of bilingual children on a Likert-scale (i.e. Not at all, Rarely, Sometimes, Often, All the time). The list of potential challenges was derived from reported assessment challenges from the literature (D'Souza et al., 2012; De Lamo White & Jin, 2011) combined with the assessment challenges faced by the researcher.

3.2.3.5 Development of Local Assessment Measures

SLPs were asked to select the type of assessment approaches, if locally developed, that could assist them in assessing Singaporean bilingual children and making accurate diagnoses. They were also asked to select the factors (i.e. test design, scoring etc.) that they thought needed to be considered during the development of local assessment measures.

3.2.4 Recruitment

SLPs who worked in paediatric settings in Singapore were invited to participate in the survey. Invitations to participate in the survey were sent via e-mail to members of the Speech and Language Therapy Singapore through their in-house email bulletin and to the point of contact of all known paediatric SLPs practices in Singapore (i.e. hospitals, community services, and private practices – a total of 49 known practices).

The survey questionnaire was made available through a Flinders University web domain where access to participate in the survey questionnaire was protected by a token. SLPs could only access the survey questionnaire via the web-link and token provided in the email invitations. To increase and maximise response rate, a second email was sent via the same route four weeks after the first email. The online survey questionnaire was open for response for nine weeks.

3.2.5 Analysis

The responses were analysed using the Statistical Program for the Social Sciences (SPSS) for Windows version 23.0. Descriptive statistics (percentages, mean, and mode) were calculated and reported.

3.3 Results

In this section, the results from Phase One are presented using descriptive statistics.

A total of 27 SLPs who provided speech therapy services to the paediatric population in Singapore participated anonymously in the survey. After analysing the reported caseload demographics for each SLP, participant #27 was not included in the final analysis as he/she reported that majority of his/her caseload consisted of non-Singaporean children who were English-speaking monolinguals. Thus, only the responses from 26 SLPs were included in the final analysis.

3.3.1 Participants and their Demographics

3.3.1.1 Total Number of Participants and Response Rate

All 26 SLPs reported that they were currently providing paediatric speech therapy services in Singapore either in the public healthcare sector (i.e. restructured hospitals: inpatients, outpatients) or VWOs (i.e. government-run therapy centres or early intervention centres) (Table 3.1). There were no exact figures for the numbers of SLPs providing paediatric speech and language therapy services in these two sectors available from the AHPC or Speech and Language Therapy Singapore to make a comparison to. Nonetheless, AHPC (2014) reported that there was a total of 237 registered SLPs working in both sectors (Table 3-1). Cruz- Ferreira and Ng (2010) previously estimated that about 49.0% of SLPs in Singapore provided paediatric services in their practice. Using the reported number and estimated percentage from both

sources, 116 SLPs (0.49 x 237) were estimated to be providing paediatric speech therapy services in both sectors.

To estimate the response rate for the survey, the number of SLPs who participated in the survey was divided by the estimated number of SLPs providing paediatric speech therapy services in both sectors (116 SLPs). The participation rate of SLPs with paediatric caseloads from these two sectors was thus estimated to be approximately 22.0% *(26/116 x 100%)*.

3.3.1.2 Demographic Data (Citizenship, Education and Experience)

The main demographic data (i.e. citizenship, country of SLP qualification etc.) of the 26 SLPs matched to the demographic data of SLPs in Singapore provided in 2014 by the AHPC (Table 3-1). Specifically, the majority of the SLPs (*n*=26; 77.0%) who participated in the survey were Singaporean citizens or Singaporean Permanent Residents, and this was comparable to data published by AHPC (72.0%) in 2014. The majority of the participants (*n*=21, 80.8%) received their SLP qualification overseas, and this was also comparable to data from AHPC (84.6%) in 2014. Most participants (*n*=17, 65.4%) had between one to six years of working experience as a SLP in Singapore which reflected the relatively young and growing SLP profession landscape in Singapore. In addition, the majority of the SLPs (*n*=21, 80.8%) who participated in the survey were currently working in Public Health Sector Services. Nonetheless, slightly more than half had previous working experience in other practice settings (i.e. VWOs, private clinics).

Table 3-1. Summary of participants' demographic characteristics compared to AHPC's data.

Characteristics	**Survey study** ***n (%)***	**AHPC, 2014** ***n* (%)**
Citizenship		
Singaporean/Singaporean Permanent Residents	20 (77.0)	259 (72.0)
Expatriate	5 (19.2)	99 (28.0)
Not stated	1 (3.8)	-
Years of experience as a SLP in Singapore		
<1 year	4 (15.4)	NDA
1-3 years	9 (34.6)	NDA
4-6 years	8 (30.8)	NDA
7-10 years	4 (15.4)	NDA
11-15 years	0 (0.0)	NDA
>15 years	0 (0.0)	NDA
Not stated	1 (3.8)	NDA
Current practice setting		
Public Healthcare Sector (restructured hospitals)	21 (80.8)	173 (48.3)
VWOs	5 (19.2)	64 (17.9)
Community hospitals	0 (0.0)	19 (5.3)
Private schools, clinics, hospitals, centres	0 (0.0)	96 (26.8)
Educational institutes	0 (0.0)	1 (0.3)
Others	0 (0.0)	5 (1.4)
Previous practice settings (if any)*		
Public Healthcare Sector (restructured hospitals)	19 (73.1)	NDA
Private hospital/clinic/school	5 (19.2)	NDA
VWOs	10 (38.5)	NDA
Others	0 (3.8)	NDA
Country of SLP qualification		
Australia	12 (46.2)	153 (42.7)
India	1 (3.8)	39 (10.9)
Ireland	4 (15.4)	23 (6.4)
Singapore	5 (19.2)	55 (15.4)
UK	1 (3.85)	35 (9.8)
USA	0 (0.0)	14 (3.9)
Others (New Zealand, Canada, not stated etc)	3 (11.5)	39 (10.9)

*Some participants have worked in at least one or more other settings previously.
NDA: No Data Available

3.3.1.3 Languages Spoken

Twenty-two (84.6%) of the 26 SLPS reported that they spoke at least two or more languages proficiently or functionally. Of the four official languages spoken in Singapore, only the English language (*n*=26, 100%) was spoken proficiently or functionally by all, followed by the Mandarin language (*n*=18; 69.2%). Despite being two of the official languages spoken in Singapore, the Malay language (*n*=1, 3.8%) and Tamil language (*n*=1, 3.8%) were reported to be spoken proficiently or functionally only by two SLPs. There were also three SLPs who each reported to be functional/proficient in the following languages: Tagalog (*n*=1, 3.8%) Hindi (*n*=1, 3.8%) and French (*n*=1, 3.8%).

3.3.2 Caseload Demographics

All 26 SLPs reported that children aged 12 and below made up 70.0% to 100.0% of the total number of clients seen on their caseload. All SLPs also reported that they assess and provide intervention to bilingual children on their caseloads who were dominant in the different languages spoken in Singapore (Table 3-2). Regardless of the languages the SLPs were able to speak proficiently or functionally, almost all SLPs (*n*=25-26, 96,2% - 100.0%) had children on their caseloads who were dominant in the English, Mandarin or Malay language (Table 3-2).

Table 3-2. Language profiles of SLP participants and language demographics of Singaporean children on their caseload.

Language	**SLPs' proficiency (proficient / functional) in:** ***n* (%)**	**SLPs who see children on their caseload that were dominant in:** ***n* (%)**	**Percentage average of children (Age 12 and below) on SLPs' caseloads who were dominant in:**	**Percentage (Singapore Department of Statistics, 2010) of Singaporean children (ages 5-9) who are dominant in:**
English*	26 (100.0)	26 (100.00)	54.3	50.5
Mandarin*	18 (69.2)	25 (96.2)	29.2	28.3
Malay*	1 (3.8)	25 (96.2)	11.9	13.1
Tamil*	1 (3.8)	19 (73.0)	3.9	4.1
Others	3 (11.4)	6 (23.1)	<1.0	NDA

*Official languages spoken in Singapore; NDA: No Data Available

The SLPs were also asked to estimate the percentage of children who were dominant in each of the official languages spoken in Singapore: English, Mandarin, Malay and Tamil. On average, about half the children on their caseload were dominant in the English language whereas the other half of children were dominant in the other official languages spoken in Singapore. The proportion of children dominant in the respective languages on the SLPs' caseloads corresponded to the census data on the language demographics of Singaporean children (Singapore Statistics Board, 2010; Table 3-2).

3.3.3 Assessment Approaches Used

Due to the diverse language landscape in Singapore, SLPs were asked to differentiate the assessment approaches used with ED and NED children to identify whether the variability in language experience among Singaporean bilingual children affected the SLPs' choices on assessment approaches used.

Overall, standardised assessments (all the time/often with NED children: 73.1%, all the time/often with ED children: 92.3%) were used most frequently (Table 3-3). This was followed by criterion-referenced measures (all the time/often with NED children: 53.8%, all the time/often with ED children: 65.4%), language sampling (all the time/often with NED children: 42.3%, all the time/often with ED children: 65.4%), DA (all the time/often with ED children: 38.5%; all the time/often with NED children 46.1%) and processing-dependent measures (all the time/often with both ED and NED children: 7.7%). Interestingly, alternative assessment approaches except for DA and processing-dependent measures were observed to be used slightly less frequently with NED children compared to ED children.

Table 3-3. Comparison of the types of assessments used by SLPs when assessing ED and NED Singaporean bilingual children.

	Assessing the language skills of Singaporean children who are ED			**Assessing the language skills of Singaporean children who are NED**		
Frequency	**All the time/Often** ***n*** **(%)**	**Sometimes** ***n*** **(%)**	**Rarely/ Not at all** ***n*** **(%)**	**All the time/Often** ***n*** **(%)**	**Sometimes** ***n*** **(%)**	**Rarely/ Not at all** ***n*** **(%)**
Standardised Assessments	24 (92.3)	2 (7.7)	0 (0.0)	19 (73.1)	7 (26.9)	0 (0.0)
Language Sampling	17 (65.4)	7 (26.9)	2 (7.7)	11 (42.3)*	9 (34.6)*	6 (23.1)*
Criterion-referenced Measures	16 (61.6)	6 (23.1)	4 (15.3)	14 (53.8)	6 (23.1)	6 (23.1)
Dynamic Assessments	10 (38.5)	5 (19.2)	11 (42.3)	12 (46.1)	4 (15.4)	10 (38.5)
Processing-dependent Measures	2 (7.7)	1 (3.8)	23 (88.5)	2 (7.7)	1 (3.8)	23 (88.5)

*Language samples in both languages.

Commercially-available standardised assessments that were created and normed for use with the UK or USA populations were used more frequently (all the time/often) than local standardised assessments (Table 3-4). Standardised assessments such the Clinical Evaluation of Language Fundamentals USA/UK Editions (e.g. *CELF 4:* Semel, Wiig & Secord, 2003) and Clinical Evaluation of Language Fundamentals Preschool USA/UK editions (e.g. *CELF P2*: Semel, Wiig & Secord, 2004) were popular choices with more than half of the SLPs reported using them frequently (all the time/often: 57.7% - 69.2%).

Out of the three available local standardised assessment, only the Singapore English Action Picture Test (SEAPT, Brebner, 2002) was reported to be used by at least half of the SLPs frequently (all the time/often: 50.0%).

Table 3-4. Standardised assessments used by SLPs when assessing the language skills of Singaporean bilingual children.

Standardised Assessments	**Frequency *n* (%)**				
	All the time	**Often**	**Sometimes**	**Rarely**	**Not at all**
***Clinical Evaluation of Language Fundamentals USA/UK Editions [CELF]* (e.g. *CELF 4:* Semel, Wiig & Secord, 2003)**	9 (34.6)	9 (34.6)	6 (23.1)	2 (7.7)	0 (0.0)
***Clinical Evaluation of Language Fundamentals Preschool USA/UK Editions [CELF P]* (e.g. *CELF P2*: Semel, Wiig & Secord, 2004)**	7 (26.9)	8 (30.8)	8 (30.8)	2 (7.7)	1 (3.8)
****Singapore English Action Picture Test [SEAPT]* (Brebner, 2002)**	5 (19.2)	8 (30.8)	9 (34.6)	3 (11.5)	1 (3.8)
***Preschool Language Scales USA/UK Editions [PLS]* (e.g. *PLS 4*: Zimmerman, Pond & Steiner, 2002)**	5 (19.2)	7 (26.9)	8 (30.8)	4 (15.4)	2 (7.7)
***Peabody Picture Vocabulary Test Editions [PPVT]* (e.g. *PVVT 4*: Dunn & Dunn, 2007)**	0 (0.0)	2 (7.7)	4 (15.4)	9 (34.6)	11 (42.3)
***Expressive Vocabulary Test Editions [EVT]* (e.g. EVT2: Williams, 2007)**	0 (0.0)	2 (7.7)	4 (15.4)	3 (11.5)	17 (65.4)
***Comprehensive Assessment of Spoken Language [CASL]* (Carrow-Woolfolk, 1999)**	0 (0.0)	1 (3.8)	6 (23.1)	6 (23.1)	13 (50.0)
****Bilingual Language Assessment Battery [BLAB]* (Sze & Rickard-Liow, 2009)**	0 (0.0)	1 (3.8)	0 (0.0)	2 (7.7)	23 (88.5)
****Cognitive Linguistic Assessment Profile [CLAP]* (Lee & Rickard-Liow, 2013)**	0 (0.0)	0 (0.0)	1 (3.8)	0 (0.0)	25 (96.2)

*Local standardised assessments.

For standardised assessments that were not created and normed for use with the Singaporean population, the majority of the SLPs (*n*=23, 84.4%) reported that they would compare the Singaporean children's performances to UK and USA normative data available in the respective manuals (Table 3-5). Other methods such as estimating the Singaporean child's

performance based on clinical experience or estimating the Singaporean child's performance descriptively by using a task analysis checklist approach were also common (>50.0%).

Table 3-5. Methods used to interpret the performance of Singaporean bilingual children on standardized assessments that were not created and normed for the Singaporean population.

Methods	**Frequency *n* (%)**		
	All the time/ Often	**Sometimes**	**Rarely/ Not at all**
Compare a child's performance to normative data provided in assessment manuals	23 (88.4)	3 (11.5)	0 (0.0)
Estimate a child's performance based on clinical experience	21 (80.8)	2 (7.7)	3 (11.5)
Estimate a child's performance using a task analysis checklist approach	13 (50.0)	8 (30.8)	5 (19.2)
Assume that Singaporean children's average language skills are about 6 months behind	8 (30.7)	5 (19.2)	13 (50.0)

3.3.4 Assessment Challenges

SLPs were asked to rate the frequency they encountered each assessment challenge when assessing Singaporean bilingual children. Table 3-6 shows that eight out of the 14 listed challenges were frequently faced (all the time, often) by the majority of the SLPs (>*n*=21,>80.8%) in their everyday clinical practice. These challenges were associated with: 1) the lack of appropriate assessment measures, 2) the lack of local normative and language development data on the languages spoken in Singapore, and 3) the lack of information on the characteristics of LI among Singaporean bilingual children.

Table 3-6. Challenges in assessment practice when assessing Singaporean bilingual children.

Challenges faced	**Frequency: All the time/ Often *n* (%)**
Availability of 'local' normative data (on standardised assessments)	24 (92.3)
Availability of information on characteristics of language impairment in Singaporean children	24 (92.3)
Availability of local developmental/normative data for the Mandarin language	24 (92.3)
Presence of linguistic bias in standardised assessments	24 (92.3)
Availability of local developmental/normative data for the Tamil language	23 (88.5)
Availability of local developmental/normative data for the Malay language	23 (88.5)
Presence of cultural bias in standardised assessments	23 (88.5)
Availability of locally adapted/developed assessments	21 (80.8)
Differentiating language disorder from language difference	18 (69.2)
Amount of time required to use a full battery of formal and alternative assessments	17 (65.4)
Ability to speak and assess the dominant language of the child	13 (50.0)
Access to interpreters	13 (50.0)
Access to known local standardised assessments	10 (37.8)
Access to international standardised assessments	4 (15.3)

3.3.5 Development of Local Assessment Measures

SLPs were asked to select the types of assessment approaches that they perceived would be useful for evaluating the language skills of Singaporean bilingual children if they were to be locally developed. The top three assessment approaches selected by SLPs were local standardised assessment (*n*=24, 92.3%), language sampling (*n*=21, 80.8%) and DA (*n*=13, 50.0%) (Table 3-7). SLPs were also asked to select the factors that they felt were important to be considered when developing local assessment measures. The top three factors selected were local developmental data (*n*=24, 92.3%), time needed for administrating the assessment (*n*=19, 73.0%) and the ease of scoring of the assessment (*n*=19, 73.0%) (Table 3-8).

Table 3-7. Locally-developed assessment approaches that can be helpful.

Assessment approaches	**SLPs who agree that this assessment approach is helpful for their practice *n* (%)**
Local Standardised Assessment	24 (92.3)
Language Sampling	21 (80.8)
Dynamic assessment	13 (50.0)
Criterion-referenced Measure	12 (46.2)
Processing-based Measure	0 (0.0)

Table 3-8. Assessment factors to consider when developing local assessment approaches.

Factors to consider during assessment development	**SLPs who agree that this factor must be considered during assessment development *n* (%)**
Test design	14 (53.8)
Time needed for administration	19 (73.0)
Local normative data	24 (92.3)
Culturally appropriate illustrations	15 (57.7)
Ease of scoring	19 (73.0)
High reliability and validity of results	15 (57.7)
High sensitivity and specificity	17 (65.4)

3.4 Discussion

In this section, the key findings from Phase One of the study are presented.

3.4.1 SLPs and their Caseloads' Language Match

The linguistic demographics of SLPs who participated in this study were unique as the majority of them spoke two or more languages proficiently or functionally (*n*=22; 84.6%). This was in contrast to the linguistic demographics of the SLPs in the USA (Caesar & Kohler: 6.2% of SLPs reported to speak at least two languages proficiently) and Australia (Williams & McLeod, 2012: 9.0% of SLPs reported to speak at least two languages proficiently) where the majority of the SLPs were reported to be only monolingual-English speakers.

However, despite being bilingual, the SLPs in Singapore faced similar challenges as their monolingual international counterparts when assessing bilingual children. Clinician and client

languages did not match. Although most SLPs were competent in English and one other official language (i.e. Mandarin, Malay, and Tamil), the children on their caseloads spoke a diverse range of languages and were dominant in other languages that the SLPs did not have any knowledge of. In particular, there were very few SLPs (*n*=2; 7.6%) who could speak Tamil or Malay although most of them reported that their caseloads consisted of children who were dominant in either the Tamil (*n*=19; 73.0%) or Malay language (*n*=25; 96.2%) (Table 3-2). For example, Participant #2 who spoke English and Mandarin proficiently/functionally reported that 20.0% of his/her caseload was children who were dominant in other languages (i.e. Malay, Tamil). Such a caseload pattern was reported by all of the SLPs, reflecting the linguistically diversified and heterogeneous nature of a multilingual country.

Most of the SLPs had to assess and provide intervention services to bilingual children whose dominant language did not match to any of those that they were proficient or functional in. A mismatch between SLP and client language has implications for the assessment process and delivery of intervention services. This may impact the SLP's ability to provide efficient and equitable speech therapy services to children whose dominant language did not match to theirs, given the lack of supporting data and materials.

The above finding demonstrated that although most of the SLPs were bilinguals, the diverse range of languages spoken in a predominantly bilingual and multilingual country will still result in a mismatch of languages spoken by SLPs and the children on their caseloads. Therefore, the solution of increasing the number of bilingual SLPs suggested by some researchers (D'Souza et al., 2012; Perrie & Core, 2006) to meet the demands of providing speech and language therapy services to bilingual children will not solve the challenges faced in assessing bilingual children. SLPs will continue to face challenges in assessing bilingual children especially if they have limited or no knowledge of the languages the children are dominant in. However, this can in part be overcome if appropriate and least-biased assessment approaches are used to assess

the language skills of bilingual children. For this reason, there is a need to better understand the different type of assessment approaches used by SLPs when assessing bilingual children's language skills and whether the assessment approaches used are appropriate for evaluating their language skills and identifying those at risk of LI.

3.4.2 Assessment Approaches Used and the Challenges

Commercially-available standardised assessments that were not designed and normed for the Singaporean bilingual population were used most frequently (i.e. all the time/often) compared to alternative assessment approaches to assess the language skills of ED and NED children (Table 3-3 and 3-4). Even when assessing the language skills of NED children, the use of commercially-available standardised assessments was still a top choice. This is despite the fact that most commercially-available standardised assessments were designed for use to assess the language abilities of monolingual English-speaking children only.

This finding is similar to the results reported in Caesar and Kohler's study (2007) where they found that SLPs working in USA school settings preferred to use standardised assessments over alternative assessment approaches when assessing the language skills of bilingual children. Even though the linguistic and cultural diversities in Singapore and USA are different, the prevalent use of commercially-available standardised assessments to assess the language skills of bilingual children are of a similar practice. This shows that limited use of recommended alternative assessment approaches to evaluate the language skills of bilingual children is a world-wide problem. In addition, this finding demonstrates that even though the majority of the SLPs were bilingual themselves in this survey, they still may not adopt recommended alternative assessment approaches in their practices.

The predominant use of commercially-available standardised assessments may be due to: 1) the lack of local normative and language developmental data for the local languages spoken in

Singapore, and 2) the lack of appropriate local assessment measures (Table 3-6). However, what is concerning was that when interpreting Singaporean children's performances on commercially-available standardised assessments, most SLPs would compare the children's performances to UK and USA normative data (*n*=23, 88.5%) (Table 3-5). This is despite the fact that the comparison of the bilingual child's score to monolingual norms is not clinically useful (De Lamo White & Jin, 2011; Thordardottir, 2005). Participant #14 justified his/her choice by stating that "*for primary school-age children, their language skills in English should have caught up sufficiently to allow for more accurate comparison against monolingual norms*". Moreover, 30.0% (*n*= 8) of the SLPs also reported that they would assume that Singaporean children's language skills are six months behind monolingual normative data. This suggests that the lack of local normative and language developmental data for the local languages spoken in Singapore has resulted in some SLPs making assumptions of the developmental trajectories of local languages based on monolingual norms.

Although there are three local standardised assessments that have been created and normed for use with the Singaporean population, only one (SEAPT: Brebner, 2002) was frequently used (all the time/often) by 50% (*n*= 13) of the SLPs in their practice (Table 3-4). Other local standardised assessments were seldom used as the majority of the SLPs reported that they either had no or limited access to them or were unaware of their existence. Those who had access but did not use them in their practice stated the limitations such as "*..does not provide sufficient information (i.e. only assesses vocabulary skills)*", "*lack of normative data for other ethnic groups"* and "*…time-consuming to use*".

Nonetheless, most SLPs seemed to be aware of the limitations of commercially-available standardised assessments and have reported the lack of appropriate local normative data (*n*=24, 92.3%), presence of linguistic bias (*n*=23, 88.5%) and cultural bias (*n*=24, 92.3%) as the top few assessment challenges faced (Table 3-6). To overcome the challenges associated with the

use of commercially-available standardised assessments, SLPs have also selected other interpretation methods such as using clinical judgement based on experience (*n*=21, 80.8%) and task analysis checklists (*n*=13, 50.0%) to estimate Singaporean children's performances instead (Table 3-5). However, these methods do not allow SLPs to appropriately determine whether the Singaporean child's performance is typically developing or is an indicator of LI.

3.4.3 Reasons for the Limited Use of Alternative Assessment Approaches

Many researchers (e.g. De Lamo & White, 2011; Hemsley et al., 2014; Laing & Kamhi, 2003) have recommended the use of alternative assessment approaches (i.e. DA, language sampling etc.) when assessing the language skills of bilingual children. Recent surveys conducted by Williams and McLeod (2012) in Australia and D'Souza and colleagues (2012) in Canada found that SLPs were using alternative assessment approaches more frequently than standardised assessments when assessing the languages skills of bilingual children. However, this was not the case in Singapore where commercially-available standardised assessments were used more frequently than alternative assessment approaches. Firstly, this could be due to the lack of local guidelines on recommended bilingual assessment practices and alternative assessment approaches. Secondly, no local study has explored the feasibility of alternative assessment approaches in evaluating the language skills of Singaporean bilingual children and whether these approaches could validly differentiate the language skills between those who are typically developing and those who are at risk of LI.

The lack of local guidelines on recommended bilingual assessment practices and alternative assessment approaches may have resulted in the reduced awareness of alternative assessment approaches among SLPs in Singapore. In countries like Australia and Canada, position papers and guidelines on bilingual assessment practices (Speech-Language and Audiology Canada, 1997; Speech Pathology Australia, 2016) have been published to create awareness and understanding among SLPs in providing equitable speech and language therapy services to

bilingual children in the respective countries. The increase in awareness on recommended bilingual assessment practices with emphasis in using alternative assessment approaches has informed SLPs in Australia and Canada to select alternative assessment approaches over standardised assessments when assessing the languages skills of bilingual children.

DA is one of the alternative assessment approaches that is gaining increasing attention in the literature as it is able to evaluate the language learning potentials of bilingual children and identify those at risk of LI with least bias (Ebert & Kohnert, 2016). However, it was one of the least used alternative assessment approaches used by SLPs in Singapore. This was similar to the results reported by Caesar and Kohler (2007) where none of the SLPs in the USA listed DA as a choice assessment method when assessing the language skills of bilingual children. Similarly, in Williams and McLeod's (2012) study, less than 20.0% of SLPs in Australia indicated the consistent use of DA when assessing the language skills of bilingual children. A possible reason is that DA is a relatively newer alternative assessment approach in the field of language testing compared to other well-established alternative assessment approaches such as language sampling and criterion-referenced assessment. Therefore, SLPs may be unfamiliar with the DA paradigm (i.e. 'Test-Teach-Retest' format) and its instructional elements (i.e. Mediated Learning Experiences: Feuerstein et al., 1979; Lidz, 1991 or Graduated Promoting: Campione and Brown, 1987). However, if there are local practice guidelines on how to adopt the use of alternative assessment approaches such as DA, it can not only increase the awareness of DA methods among SLPs but also increase their familiarity and confidence in conducting DA in their practice.

Secondly, besides the lack of practice guidelines in Singapore, no local study has investigated the use of alternative assessment approaches to differentiate the language performances between typically developing Singaporean bilingual children and Singaporean bilingual children with LI. As a result, SLPs in Singapore may not be confident in adopting alternative

assessment approaches in assessing Singaporean bilingual children and tended to over-rely on commercially-available standardised assessments as they are known to be "tools of the profession" (Stow & Dodd, 2003. P. 363).

3.4.4 Development of Local Assessment Measures

The over-reliance on commercially-available standardised assessments points to a need for developing local assessment measures, in particular alternative assessment measures, that SLPs in Singapore can use to validly assess the language skills of Singaporean bilingual children. Although there are currently three locally developed standardised assessments, they were not used as an assessment of choice. In fact, two of the assessments were not used at all by the majority of the SLPs (>88.5%). SLPs reported reasons such as: 1) these locally developed standardised assessments assessed only limited language areas, 2) there was a lack of local normative data for other ethnic groups, and 3) the assessments had lengthy and time-consuming assessment procedures. Therefore, before any more local assessment measures are developed for use with the Singaporean bilingual population, it is essential to understand: 1) the types of assessment approaches SLPs in Singapore will find useful in their practice and, 2) the assessment factors that need to be considered when developing local assessment measures.

In the survey, SLPs were asked to identify the language assessment approaches that they perceived useful in their practice if they were to be locally developed (Table 3-7). Although most SLPs had selected local standardised assessment (*n*=24, 92.3%) as a top choice, the development of local standardised assessments with local normative data for bilingual populations is highly challenging (Erbet & Kohnert, 2016). The heterogeneous nature of the Singaporean bilingual population where four official languages are commonly spoken makes it almost impossible to gather normative data for the population. Brebner and colleagues (2016) have also previously highlighted that the development of local normative data for the Singaporean population is complex due to the variability in language experience and language

dominance. Even if normative data was gathered, the patterns of language use within the child and within the Singaporean community are changing over time, impacting the validity and applicability of such data. Therefore, there is a need to explore the use of alternative assessment approaches instead to evaluate the language skills of Singaporean bilingual children.

Among the list of alternative assessment approaches, the top two alternative language approaches selected by SLPs were language sampling (*n*=21, 80.8%) and DA (*n*=13, 50.0%). However, as highlighted previously (Section 2.4.3), language sampling requires the comparison of the child's language sample to local normative or language developmental data of the languages spoken in the community. The developmental trajectories of the local languages spoken in Singapore have not been thoroughly studied. Until sufficient local normative and language developmental data have been gathered and published, language sampling may not be an appropriate alternative assessment approach to be adopted in local practices. DA, on the other hand, does not require the comparison to local normative or language developmental data, to accurately identify bilingual children at risk of LI. DA is, therefore, a potential alternative assessment approach that can be explored and developed in the near future to assist SLPs in evaluating the language learning skill of Singaporean bilingual children and therefore assist in identifying those at risk of LI.

SLPs were also asked to select the factors that were deemed important to be considered when developing local assessment tools (Table 3-8). The availability of 'Local normative data' (*n*=24, 92.3%) was a top factor but is only applicable if local standardised assessments are to be developed. 'Time needed for administering the assessment' (*n*=19, 73.0%) and 'Ease of scoring' (*n*=19, 73.0%) were selected as the next top two factors to be considered. SLPs want language assessments that are relatively quick to administer and easy to score while providing useful diagnostic information about the child. For example, two out of the three locally-developed standardised assessments were not used at all as SLPs shared that they were *"time-consuming*

to use" and "*..does not provide sufficient information".* Thus, to increase the clinical acceptability of new language measures that are locally developed, it is important to ensure that the measures can be administered and scored within a short period of time while providing important clinical information that can assist SLPs in differential diagnosis.

3.4.5 Limitations and Future research

The results obtained from the survey study provide insights into the assessment practices and challenges faced by bilingual SLPs working with a predominantly bilingual population in Singapore. However, due to the small sample size (*n*=26), and the exploration of the experiences in only one predominantly bilingual country, the results cannot be generalised to the assessment practices of all SLPs in Singapore or around the world. The SLPs who participated in the study were all currently working in the Public Healthcare Sector. The type of services provided in the Public Healthcare Sector may have accounted for their choice of assessment. However this was not explored in the study as they were not asked to list the reasons why they chose certain assessment approaches over others. As such there could be other factors which were not explored (e.g. modes of service delivery in Singapore) that could have affected SLPs' choice of assessments. This is one area to investigate in the future to understand why SLPs in Singapore felt compelled to use standardised assessment over alternative assessments.

Future research can investigate the assessment practices of SLPs working in other predominantly bilingual countries and compare the assessment practices and challenges faced. In countries where bilingualism is the norm, alternative assessment approaches such as DA can be developed and explored to evaluate the language skills of bilingual children. This may consequently lead to an increase in awareness and adoption of alternative assessment approaches in clinical practice. Other potentially useful research could be to collect local normative and language developmental data for the languages spoken in the population and/or

identify languages tasks/markers that can successfully differentiate language disorder from language difference among bilingual children.

3.4.6 Conclusion

Bilingualism has increasingly become the norm in the world of today. Some have suggested increasing the number of bilingual SLPs so as to meet the increasing demand to provide SLP services in linguistically diverse (bilingual) communities (e.g. D'Souza et al., 2012; Perrie & Core, 2006). Findings from Phase One show that bilingual SLPs also face similar challenges as their monolingual international counterparts in assessing bilingual children from culturally and linguistically diverse communities. The solution of increasing the number of bilingual SLPs globally alone will not resolve the challenges faced in assessing bilingual children.

Although commercially-available standardised assessments are not recommended to be used to assess the language skills of bilingual children, there is still an over-reliance on commercially-available standardised assessments to assess bilingual children's language skills in predominantly bilingual countries like Singapore. This can be attributed to 1) little or no information on local normative and language developmental data for the languages spoken, 2) lack of local practice guidelines on recommended bilingual assessment practices and alternative assessment approaches and, 3) lack of local studies exploring the feasibility and utility of alternative assessment approaches in assessing the language skills of bilingual children and therefore identifying those at risk of LI.

Recent studies (Camilleri & Law, 2007; Hasson, Camilleri, Jones, Smith & Dodd, 2013) found that alternative assessment approaches such as DA can accurately identify LI from language difference among bilingual children compared to commercially-available standardised assessments. However, findings from Phase One found that alternative assessment approaches are still not widely adopted or accepted even in a predominantly bilingual country like

Singapore where bilingual children make up the majority of SLPs' caseloads. Globally, there is an ongoing need to translate such research outcomes into actual clinical practice, especially in countries/communities where bilingualism is the norm. To do so and as part of evidence-based practices, it may be worthwhile to conduct studies in countries/communities (i.e. Singapore) where bilingualism is the norm to explore the use of alternative assessment approaches to evaluate the language skills of bilingual children and therefore identify those at risk of LI. With the information and evidence from such studies, SLPs working in predominantly bilingual countries may then be confident in adopting alternative assessment approaches into their clinical practices.

3.5 Chapter Summary

In the first section (Section 3.2) of the chapter, the local context in which Phase One of the study was conducted was elaborated upon. The survey study was conducted in Singapore as it is a predominantly bilingual country. The unique language landscape in Singapore provides an ideal setting to investigate the assessment practices and challenges faced by SLPs working in a predominantly bilingual country. A survey questionnaire was developed, and SLPs providing paediatric speech therapy services in Singapore were invited to participate. A total of 26 responses were obtained and analysed.

In the second (Section 3.3) and third section (Section 3.4) of the chapter, the results and discussion of the survey study were presented. The key findings of the survey study were as follows:

1. Despite being bilingual themselves, the SLPs in Singapore continued to face multiple assessment challenges that were related to:
    a. Little or no information on local normative and language developmental data for the languages spoken in Singapore.

b. Lack of appropriate local assessment approaches. No local practice guidelines on the recommend assessment practices and alternative assessment approaches.
c. Lack of local studies exploring the feasibility and utility of alternative assessment approaches in identifying LI among Singaporean children.

2. Internationally, SLPs are still relying on commercially-available standardised assessments that were designed for use with monolingual children to assess bilingual children. Although there has been an increased awareness of alternative assessment approaches to assess the language skills of bilingual children, little has been translated into clinical practice.
3. To contribute to the landscape on recommended bilingual assessment practices especially in countries where bilingualism is the norm, there is a need to explore and develop the use of alternative assessment approaches to evaluate the language skills of bilingual children to determine whether alternative assessment approaches can assist SLPs in identifying those at risk of LI.

Chapter 4 Phase Two: Literature Review Part A

4.1 Introduction

Alternative assessment approaches are highly recommended over commercially-available standardised assessments for assessing the language skills of bilingual children and identifying those at risk of LI (De Lamo White & Jin 2011; Laing & Kamhi, 2003). However, one of the main findings from Phase One was that even in a predominantly bilingual country like Singapore where the majority of the caseloads were bilingual children, there was still an over-reliance on commercially-available standardised assessments that were designed and normed for use with monolingual children. Therefore, to increase the awareness and adoption of alternative assessment approaches when assessing the language skills of bilingual children in countries where bilingualism is the norm, there is a need to conduct local studies to explore the use of alternative assessment approaches to assess and identify bilingual children at risk of LI. In particular, whether alternative assessment approaches like DA can better identify bilingual children at risk of LI compared to commercially-available standardised assessments that are commonly used.

A number of alternative assessment approaches have been previously discussed (Chapter Two, Section 2.4.2 to Section 2.4.5). However, many of them (i.e. language sampling, criterion-referenced measures) require comparison to local normative or language developmental data for the languages spoken in the bilingual community in order to differentiate LI from language difference among bilingual children. The reality is that the availability of local normative data and language developmental data for the languages spoken in most bilingual communities/countries is often limited or unavailable due to the lack of research. There is, however, one recommended alternative assessment approach that does not require the use of or comparison to local normative data and language developmental data: dynamic assessment (DA). Unlike other alternative assessment approaches where children's language ability is

assessed by evaluating existing language knowledge, DA assesses children's language ability by evaluating their potential to learn new language materials (i.e. language learning potential).

The primary objective of Phase Two of the study is to develop and gather validity evidence on whether a DA process can validly differentiate the language learning potential between bilingual children with and without LI, and therefore accurately identify those at risk of LI. To do so, a chain of evidence under the Validity Argument Framework (Kane, 2006; 2013) is gathered to support the proposed interpretations and uses of the DA process. The intention is not to develop a finalised assessment tool but to explore and provide preliminary validity evidence that DA can evaluate the language skills of young bilingual children with least bias and with more accuracy compared to a commercially-available standardised assessment. It is hoped that the evidence gathered can increase the awareness and clinical acceptability of DA as a viable language assessment approach to evaluate the language learning potential of bilingual children and therefore identify those at risk of LI in predominantly bilingual countries.

In this chapter, the origin and concept of DA are further outlined. The differences between DA and familiar static assessment approaches (i.e. standardised assessments etc.) will be highlighted. Next, the two approaches that are commonly used in DA to evaluate language learning potential, Mediated Learning Experiences (MLE; Feuerstein et al., 1979; Lidz, 1991) and Graduated Prompting (GP; Campione and Brown; 1987) will be discussed. The principles, advantages and potential limitations of each approach will be elaborated upon. At the same time, the research for each DA approach in discriminating LI from typical development will be presented. This is followed by the identification of potential research gaps on DA approaches in the field of language testing which will form the basis of the relevance and purpose of Phase Two of the study. Finally, the importance of gathering and evaluating a chain of validity evidence using the Validity Argument Framework (Kane, 2006; 2013) to support the

interpretation and use of a proposed DA process in discriminating between LI and typical development among young bilingual children is highlighted.

4.2 Dynamic Assessment: Evaluation of Language Learning Potential

Before presenting the origin and concept of DA, it is important to make the distinction between DA and static assessment in the field of language testing. Static assessments are assessment approaches where children's language abilities are evaluated based on their existing language knowledge in specific language domains (De Lamo White & Jin, 2011; Laing & Kamhi, 2003). Static assessments focus on the identification of LI by evaluating children's current language performance in specific language task/s in comparison to their same-aged peers. Assessment approaches such as standardised assessments and criterion-referenced measures are some examples of static assessments. DAs, in contrast to static assessments, evaluate children's ability to acquire new language knowledge within the assessment itself.

Bilingual children have variable language and learning experiences. Therefore, single 'snapshots' of their language knowledge on static assessments may not provide accurate evaluations of their actual language ability (De Lamo White & Jin, 2011). The evaluation of bilingual children's language skills on static assessments can be biased. A bilingual child's poor performance on static assessment may not necessarily be an indication of LI but may be due to the difference in language experience or opportunities (i.e. language difference). This can lead to an over-identification of LI among bilingual children when they, in fact, have normal language ability. Conversely, this can also lead to an under-identification of LI among bilingual children especially if SLPs attribute poor performance to a language difference (De Lamo White & Jin, 2011; Laing & Kamhi, 2003).

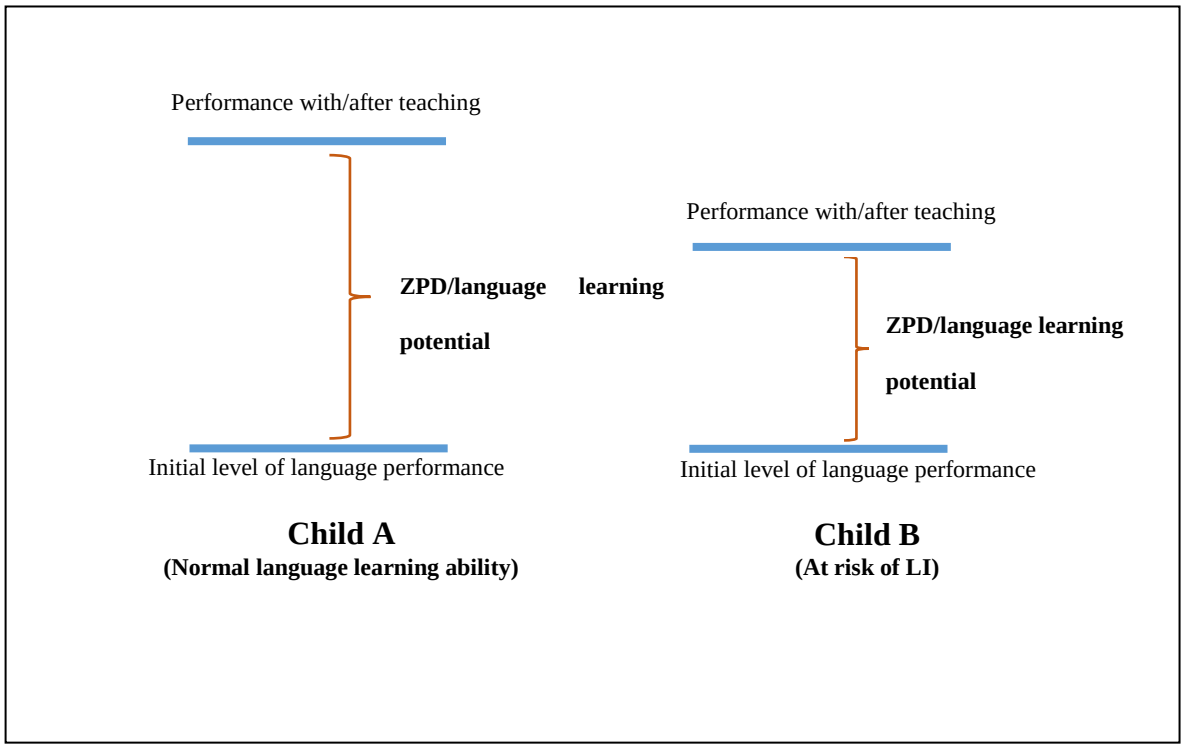
DA is a recommended alternative assessment approach for assessing bilingual children's language skills compared to static assessments as it reduces the assessment bias associated with

previous language experiences (Gutiérrez-Clellen & Peña, 2001; Hasson et al., 2012; Kapantzoglou et al., 2012; Petersen, Chanthongthip, Ukrainetz, Spencer & Steeve, 2017). Unlike static assessments where children's language abilities are determined by their existing language knowledge with comparison to normative data or a set criteria, DAs evaluate children's ability to acquire new language knowledge or skills during the assessment itself. DA has been shown to be able to differentiate between LI and language difference among bilingual children, therefore reducing the risk of over or under-identification of LI among the said population (Camilleri & Law, 2007).

DA is conceptualised from Vygotsky's (1978) work on sociocultural theory and cognitive development theory. Vygotsky (1978) first recognised that children from 'disadvantaged backgrounds' (i.e. bilingual or culturally and linguistically diverse background, lack of stimulation, immigrant children) were often mistakenly identified as having poor cognitive levels on static cognitive assessments (i.e. standardised cognitive assessments). Vygotsky believed that cognitive development in children (i.e. language learning, information processing etc.) occurs and develops through socialization opportunities and interactions in the environment before being internalised as higher cognitive functions (i.e. concept formation problem-solving, working memory etc.). Therefore, he proposed the concept of Zone of Proximal Development (ZPD). ZPD is defined as the distance between the "actual developmental level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). This is also known as learning potential (Lidz, 1991). Vygotsky believed that evaluating the learning potential of a child was a better indicator of the child's actual cognitive ability and future prognosis. Although he did not provide details of a DA format, he laid the groundwork for the development of DA procedures by later researchers.

In the field of language testing, the goal of DA is to determine the language learning potential of the child by determining the 'width' of the ZPD through the child's interaction with an examiner (Gutiérrez-Clellen & Peña, 2001). As illustrated in Figure 4-1 below, two children with the same initial level of language performance can have different widths of ZPD. Child A has a 'wide' ZPD which indicates high language learning potential. Child A is expected to be able to learn independently in a socially supportive environment (i.e. classroom). Child A is likely a child with normal language learning ability. Child B, however, has a 'narrow' ZPD which indicates lower language learning potential. Child B will likely struggle to learn independently even in a socially supportive learning environment and will require intervention or/and learning support. Child B is likely a child at risk of LI and will require further assessment and intervention.

Figure 4-1. An illustrative example of ZPD.



In the following section, the common DA approaches and procedures used in the field of language testing will be elaborated upon.

4.3 Dynamic Assessment Approaches in Language Assessment

Two main formats of DA have been developed based on Vygotsky's concept of ZPD – the 'Test-Teach-Retest' format and 'Testing the Limit' format. In the 'Test-Teach-Retest' format, children take a pre-test (i.e. Test phase) that evaluates their current skills in a specific task. Next, in the 'Teach' phase, they are provided with teaching instruction by the examiner in the skills measured during the pre-test. After teaching is completed, the children's skills are re-evaluated in a post-test (i.e. Re-test phase). The post-test may be the exact same test used as during pre-test or an alternative but parallel form of the pre-test (Sternberg and Grigorenko, 2002). In the 'Testing the Limit' format, testing and teaching instructions are provided for each test item consecutively. For example, the child's skill is assessed on the first test item. If he or she answers or solves it correctly, the next test item will be presented. However, if the child does not respond correctly, teaching instructions are provided by the examiner until the child is able to respond to the test item correctly (Sternberg and Grigorenko, 2002).

The 'Test-Teach-Retest' DA format is commonly adopted in the field of language testing to differentiate LI from language difference among children from bilingual backgrounds (Gutiérrez-Clellen & Peña, 2001). The 'Test' phase functions as a baseline where children's language knowledge in a specific language task is evaluated. This is followed by the 'Teach' phase where teaching instructions are provided by the examiner (i.e. SLP) to learn the new language materials or skills required to complete the language task assessed in the 'Test' phase. The teaching instructions provided can vary, some are unstructured or less structured, such as Mediation Learning Experience (MLE: Feuerstein et al., 1979; Lidz, 1991), and others are structured such as Graduated Prompting (GP: Campione and Brown; 1987). The 'Retest' phase then evaluates the learning and maintenance of the language materials or skills addressed in the 'Teach' phase.

For DAs that incorporate MLE principles as teaching instructions, the examiner observes how the child learns, detects the difficulties faced in the learning process and provides individualised teaching instructions (i.e. mediation) during the 'Teach'phase (Kapantzoglou et al., 2012). The child's response and learning progress serve as feedback for the examiner to adjust the amount of mediation to be provided. The teaching instructions provided are based on Feuerstein's mediation principles (i.e. *Intentionality, Transcendence, Meaning, and Competence*: Feuerstein et al., 1979; Lidz, 1991). The function of mediation is to provide the child with cognitive tools that will enable the child to learn. These principles will be further elaborated upon in the subsection (Section 4.3.1) below. On the other hand, for DAs that incorporate GP principles as teaching instructions, the examiner observes how the child learns and provides pre-determined teaching instructions (i.e. cues) during the 'Teach phase'. The teaching instructions are provided in a hierarchical order based on the child's learning responses. The GP method will be further elaborated upon in the subsection (Section 4.3.2) below.

The final outcome of DA is to determine the child's language learning potential. Each child's learning potential is determined in two ways. Firstly, by evaluating the child's responses to the teaching instructions provided in the 'Teach phase', also known as *modifiability* (Feuerstein et al., 1979). Secondly, by evaluating the gains in the child's language skill/performance in the 'Retest' phase (i.e. improvement in test scores from 'Test' phase to 'Retest' phase if the same test is used). A child that demonstrates a significant change in language skill/performance (i.e. a significant improvement in scores in 'Retest' phase) and high modifiability (i.e. responsive to teaching instructions, minimum examiner's effort is required) is likely a typically developing child with good language learning potential. On the other hand, a child who demonstrates little change in performance (i.e. a poor gain in scores in 'Retest' phase) and low modifiability (i.e.

unresponsive to teaching instructions, maximum examiner's effort is required) is likely to be a child with poor language learning potential. This child is likely to be at risk of LI.

In the next two subsections, the key features of the MLE approach to DA (Section 4.3.1) and the GP approach to DA (Section 4.3.2) will be elaborated upon. Studies in the field of language testing that used either approach are also reviewed to highlight the advantages and limitations of each approach in the evaluation of children's language learning potential.

4.3.1 Mediated Learning Experience Approach to Dynamic Assessment

MLE is designed to teach children about the principles of task solution and problem-solving strategies to complete a task successfully (Peña et al., 2001). Feuerstein and colleagues (1979) believed that children from disadvantaged backgrounds (i.e. low exposure to the language being assessed, low social economic status) could present with inefficient cognitive/information processing skills in a specific domain (i.e. a language task) due to the lack of relevant experience or a lack of stimulation. Therefore, MLE can help these children, who were previously not exposed to the language task, to develop and use cognitive tools (i.e. classification, comparison, analytical perception, analogical reasoning) to learn and complete the language task successfully (Peña, Reséndiz & Gillam, 2007). The ease with which children can learn and employ cognitive-linguistic skills is an indication of the child's modifiability (Peña et al., 2007). Children with LI will have difficulties in learning and applying cognitive tools, indicating poor modifiability (Botting & Conti-Ramsden, 2001; Ellis Weismer, Evans, & Hesketh, 1999; Gillam, Cowan, & Marler, 1998; Jordaan, Shaw-Ridley, Serfontein, Orelwitz, & Monaghan, 2001). On the other hand, when provided with cognitive tools, children from disadvantaged backgrounds but with typical language learning ability can usually demonstrate the ability to learn efficiently, indicating high modifiability.

MLE focuses on giving children the foundation for competence and prerequisite experience with a targeted language task (Peña et al., 2001). In general, there are four mediation components that are essential for an MLE: *intentionality, transcendence, meaning and competence (*Feuerstein et al., 1979; Lidz, 1991). *Intentionality* mediations refer to the examiner's intent and conscious effort in engaging the child in an interaction for the purpose of the teaching. To do so, the examiner explicitly states the goal of the teaching (i.e. learn new language materials) that is about to occur to help the child be aware of the purpose of the interaction (Lidz, 1991; Peña, 1996; Peña et al., 2001). Next, the examiner provides *transcendence* mediations to help the child to think hypothetically about the goal of the immediate task by linking it to familiar events that the child can relate to or has experienced (Lidz, 1991; Peña, 1996; Peña et al., 2001). This is immediately or concurrently followed by the examiner providing *meaning* mediations to highlight the important features of the task to be learnt. This directs the child's attention to the critical features of the task and enhances the child's awareness of what to notice and why learning the task is important (Lidz, 1991; Peña, 1996; Peña et al., 2001). Finally, the examiner provides *competence* mediations to help the child develop, plan and carry out cognitive strategies to learn and complete the task with competence (Lidz, 1991; Peña, 1996; Peña et al., 2001). The ultimate goal of MLE is to help the child become a self-regulated and active learner (Peña et al., 2001). An advantage of the MLE approach to DA is that the mediation components provide a meaningful and engaging context for targeted learning to take place. This creates a naturalistic and engaging environment for testing to take place while reducing any test anxiety that children may have.

Peña and colleagues (1992; 2001) were pioneers in the field of language testing that adopted an MLE approach to DA to evaluate the language learning potential of bilingual children with and without LI. In their studies, the examiner not only has to evaluate the gains in the child's language skill/performance in the 'Retest' phase but also assign modifiability scores to reflect

the child's responsiveness to the MLE provided and the examiner's effort in providing the required amount of MLE to support the child's learning during the 'Teach' phase. These scores, known as the Modifiability Index, are allocated subjectively by the examiners on either or both of the following Likert scales checklists: the Learning Strategies Checklist (Lidz, 1991; Peña, 1993, Peña, 2000) and the Modifiability Scale (MS: Lidz, 1991; Peña, 2000). The combination of both 'Retest' score and Modifiability Index score produced the best discrimination result. For example, Peña and colleagues (2001) investigated the word learning potential of 45 typically developing bilingual preschool children and 10 bilingual preschool children who were reported to have low-language ability (i.e. children with possible LI) on a DA of vocabulary. The Expressive One-Word Picture Vocabulary Test-Revised (Gardner; 1990) was used as a measure of vocabulary (i.e. language ability in a specific task) at the 'Test' and 'Retest' phases. The children were taught naming strategies using MLE. Results showed that the combination of 'Retest' scores and Modifiability Index achieved the best classification accuracy (i.e. typically developing as typically developing, LI as LI) with an overall correct classification accuracy of 92.3%. However, if the Expressive One-Word Picture Vocabulary Test-Revised (Gardner; 1990) was used alone to evaluate the vocabulary knowledge of the children, the overall correct classification accuracy was only 25.0%. In another study, Peña and colleagues (2014) investigated the narrative learning potential of 18 Spanish-English preschool children with LI and 18 typically developing Spanish-English preschool children on a DA of narrative ability. The story intervention script incorporated key elements principles of MLE which focused on the creation of complete and complex story episodes. Results of the study showed that the combination of Modifiability Index and 'Retest' scores also achieved a high classification accuracy of 80.6% to 97.2%.

Besides studies conducted by Peña and colleagues, other researchers who also adopted the MLE approach to DA in their studies were also able to discriminate the language learning

potential of bilingual children with LI and without LI in their respective studies with high classification accuracy. For instance, Kapantzoglou and colleagues (2012) investigated the language learning potential of 13 Spanish-English preschool children with LI and 15 Spanish-English typically developing preschool children on a DA of word learning ability. To minimise the bias of previous language experience, three novel words that followed Spanish phonological rules instead of real Spanish words were taught to participants in the 'Teach' phase in a structured play activity using an MLE approach. Results indicated that 'Retest' scores combined with Modifiability Index scores achieved a high classification accuracy for both groups with sensitivity and specificity reported at 76.9% and 80.0% respectively. In another study, Petersen and colleagues (2017) investigated the performance of 10 Spanish-English bilingual children with LI and 32 typically developing Spanish-English bilingual children aged between 6;4 to 9;6 on a DA of narratives. At the 'Test' phase, children were assessed on their ability to retell a story based on the amount of story grammar and subordinate conjunctions used ('Test' scores). During the 'Teach' phase, the examiners taught the children missing grammar elements and subordinate conjunctions using a structured MLE approach. Lastly, during the 'Retest' phase, children were asked to retell a parallel story to determine whether there was an increase in the amount of story grammar and subordinate conjunctions used ('Retest' scores). Results indicated that once again the 'Retest' scores combined with Modifiability Index scores achieved high classification accuracy for both groups with sensitivity and specificity rates over 90.0%,

Although findings from above research studies have suggested that an MLE approach to DA is a promising assessment measure for accurately discriminating the language learning potential of bilingual children with LI and typically developing bilingual children, there are administrative barriers that may hinder its clinical acceptability and adoption into clinical practice. For example:

1. The provision of MLE teaching instructions is often unscripted. SLPs have to undergo extensive training to be familiar with the implementation of MLE principles to skillfully provide mediation instructions in response to each child's response during the assessment.
2. SLPs have to be trained in the rating and allocation of Modifiability Index scores using the available Likert-scale checklists (i.e. Learning Strategies Checklist: Lidz, 1991; Peña, 1993, Peña et al., 2000; Modifiability Scale: Lidz, 1991; Peña, 2000).
3. The administration of mediation instructions for each child is individualised, and scoring of Modifiability Index by examiners is subjective (Hasson & Joffe, 2007; Patterson, Rodríguez & Dale, 2013; Petersen et al., 2017). This makes the analysis and interpretation of results seem less reliable and valid.
4. Highly individualised mediation instructions are required during the DA. As such, the DA process is thought to be effortful and time-consuming to conduct (Hasson & Joffe, 2007; Patterson et al., 2013; Petersen et al., 2017).

In summary, there is an emerging body of evidence for an MLE approach to DA to differentiate between the language learning potential of bilingual children with and without LI and therefore identifying those at risk of LI. However, the application of MLE principles in DA may be challenging for SLPs who are unfamiliar and untrained in MLE concepts, hindering its clinical acceptability and adoption into clinical practice.

4.3.2 Graduated Prompting Approach to Dynamic Assessment

GP uses a series of pre-determined prompts to progressively provide the child with more support and information to learn in the DA (Campione & Brown, 1987). The pre-determined prompts are designed to vary in the level of contextual support they provide, from least assistive to most assistive. The child's modifiability is defined by the number of pre-determined graduated prompts for the child to learn and complete the language task.

An advantage of the GP approach to DA is that the graduated prompts provided are scripted and pre-determined. Moreover, the pre-determined prompts are usually based on cueing hierarchies that SLPs are familiar with as part of the language intervention work they do with children. Therefore, examiners do not have to undergo additional training to administer DAs that incorporate GP principles as the scripts and scoring guidelines are self-explanatory and objective (i.e. evaluating the number of pre-determined prompts required).

Bain and Olswang (1995; Olswang & Bain, 1996) were the pioneers in the field of language testing to adopt a GP approach to DA by outlining a hierarchy of pre-determined prompts to evaluate children's language learning potential. For instance, in their 1995 study, Bain and Olswang evaluated the learning language potential of 15 monolingual preschool children with specific expressive LI in learning two-word utterances over a period of nine weeks (i.e. Initial assessment: 3 weeks, Treatment - DA: 3 weeks, Follow-up: 3 weeks). A set of pre-determined prompts from least assistive to most assistive (i.e. general statement, elicitation questions, sentence completion, indirect model, direct model evoking a spontaneous imitation, direct model with an elicitation statement) were provided in a hierarchical order to support the children in producing 54 sets of two-word utterances during the treatment period. Results showed that children who were more responsive to the cueing hierarchy (i.e. required less supportive prompts to produce all 54 sets of two-word utterances) during the treatment period demonstrated greater language production change throughout the nine weeks study period than children who were less responsive to the cueing hierarchy (i.e. required more supportive prompts). They concluded that the GP hierarchy possessed predictive validity and was useful in determining the children's modifiability and therefore their language learning potential.

Given the evidence that monolingual children's modifiability could be evaluated in DAs using a GP framework, researchers began to explore whether the GP framework could objectively evaluate the language learning potential of bilingual children from different backgrounds (i.e.

Burton & Watkins, 2007; Patterson et al., 2013). For example, Burton and Watkins (2007) evaluated the word learning ability of 24 typically developing African-American preschool bilingual children on a DA of word learning task versus a static assessment of vocabulary. In their study, 12 children were grouped in the high-risk background group (i.e. received free lunch in school reflecting low social-economic-status, mother's highest educational qualification was high school or below) while the other 12 children were grouped in the low-risk background group (i.e. did not receive free lunch in school reflecting average social-economic status, mother's highest education qualification was at least college and above). It was hypothesised that because all 24 children were typically developing, they would perform similarly on the DA word learning task but not on the static assessment of vocabulary as the high-risk group would have less relevant language experience/exposure. In the DA, four target words (non-words) and their matched object referents were presented to the children in a picture storybook that was narrated by the examiner. Each target word was presented eight times in the story. After the story was presented, toys that corresponded to each of the referent object used in the story were presented. The children were asked to name the toys using the target words. A set of pre-determined prompts from least assistive to most assistive (i.e. elicitation question, elicitation with reference, elicitation in cloze format, election with mode) were provided successively until the child could name the toy with its paired target word. The children's word learning abilities were evaluated by tabulating the total number of prompts required to produce the name of all four toys using the target words. Results showed that the total number of prompts required to produce all four novel words was not significantly different between the two groups as hypothesised. However, on the static assessment of vocabulary, the group of children from high-risk background obtained significantly lower scores compared to the group of children from low-risk background. The researchers concluded that using DA

with a GP framework can impartially evaluate the word learning ability of typically developing bilingual children from different backgrounds compared to a static assessment of vocabulary.

Researchers also began to evaluate whether DAs with GP framework could differentiate the language learning performance of bilingual children with and without LI (i.e. Camilleri & Law, 2007; Hasson et al., 2012). Hasson and colleagues (2012) examined the language learning potential of 12 bilingual children who were referred for speech and language therapy and 14 typically developing bilingual children on vocabulary, sentence structure and phonology tasks. In all three language tasks, the baseline of their performance in each language task was first measured in the 'Test' phase. In the 'Teach' phase, four to 10 target items were presented for each language task. A hierarchy of predetermined prompts from least assistive to most assistive was provided for the children to learn all target items in each language task successfully. The children's performance in all three language tasks was re-evaluated in the 'Retest' phase. Results showed that the performance of both groups was differentiated by the total number of prompts required in the 'Teach' phase and 'Retest' scores across all three language tasks. Overall, the referred group of bilingual children required more prompts than their typically developing counterparts in the 'Teach' phase to learn across all three language tasks. Despite receiving more prompts during the 'Teach' phase, the referred group of bilingual children still obtained lower 'Retest' scores on all three language tasks compared to their typically developing counterparts. Although classification accuracy (i.e. sensitivity and specificity) was not determined in the study, the results demonstrated that combination of the number of prompts required in the 'Teach' phase and 'Retest' phase scores differentiated the language learning potential of bilingual children with and without LI.

In summary, the evidence for a GP approach to DA to differentiate between the language learning potential between bilingual children with and without LI is emerging. However, it is unsure if the GP approach to DA is an effective assessment approach in identifying bilingual

children at risk of LI as classification accuracies were not determined in all studies. Therefore, research is needed to explore the classification accuracy of a GP approach to DA in differentiating between bilingual children with and without LI.

4.4 Dynamic Assessment: The Research Gaps

The increase in awareness of the limitations of static assessments in assessing the language skills of bilingual children has led to an increase in research on the use and exploration of DA as an alternative assessment approach for this population. However, the research has only been conducted in predominantly monolingual countries with minority groups of bilingual children (i.e. Spanish-English bilingual children in the USA, culturally and linguistically diverse groups of children in the UK). To the researcher's knowledge, the feasibility and utility of DA as an appropriate alternative assessment approach to distinguish between typically developing bilingual children and bilingual children in predominantly bilingual countries have not been explored. This may be a contributing reason for why DA is not a preferred choice of assessment in evaluating the language skills of bilingual children in predominantly bilingual countries like Singapore (Teoh et al., 2017).

To increase the awareness and adoption of DA as a preferred choice of assessment when assessing bilingual children in predominantly bilingual countries, it is crucial to conduct local studies to provide evidence that DA approaches can more accurately differentiate between bilingual children with and without LI compared to commercially-available standardised assessments. At the same time, to increase the clinical acceptability of DA as a preferred assessment approach for assessing bilingual children, it is necessary to develop and explore a DA process that is easy for SLPs to administer and to score children's modifiability and language learning potential.

A review of the research on MLE approach to DA (Section 4.3.1) has identified its potential for use in clinical settings to evaluate the language learning potential of bilingual children and therefore identify those at risk of LI. Moreover, its approach of providing a naturalistic test environment (i.e. mediation components) where a meaningful and engaging context is provided within the assessment to help children to learn will appeal to SLPs, increasing its clinical acceptability. However, there are administrative barriers as previously presented (Section 4.3.1) that may hinder the adoption of MLE approach to DA in clinical settings. To overcome some of the administrative barriers, a few studies (i.e. Kapantzoglou et al., 2012; Petersen et al., 2017) have designed the provision of MLE instructions in DAs as a script to increase the ease of administering MLE instructions. This was to assist SLPs who have not had training or are unfamiliar with the MLE approach to administer the DA process with consistency (i.e. fidelity). However, the scoring of children's modifiability in these studies was still based on the examiners' subjective ratings of the children's responsiveness to the teaching instructions on modifiability rating scale checklists. This would still require SLPs to undergo additional training to be reliable with the ratings, hindering the clinical acceptability of MLE approach to DA.

Besides the possible administrative barriers that were previously mentioned (Section 4.3.1), there are additional barriers that may hinder the adoption of MLE approach to DA in bilingual clinical settings, where young bilingual children make up the majority of the caseloads. For example:

1. The provision of MLE instructions relies heavily on language as the mode of delivery (Camilleri & Bottling, 2013). Bilingual children often present with variable language experiences. Bilingual children whose dominant language does not match to the language in which the MLE instructions are provided may have difficulties following and understanding the MLE instructions provided.

2. The metacognitive skills that are required to follow and understand MLE type of instructions are still developing in young bilingual children (i.e. preschool children). Therefore, they may not fully benefit from the provision of MLE instructions and may face difficulties in applying the cognitive tools provided through MLE instructions to complete the targeted language tasks in the DA (Haywood & Lidz, 2006).

On the other hand, as compared to the MLE approach to DA, the GP approach to DA is comparatively more clinically 'user friendly'. This is because the GP approach is highly scripted as the hierarchy of graduated prompts provided is always pre-determined. Moreover, the scoring of the children's modifiability is objective as it is based on the number of graduated prompts provided by the SLPs. Although a review of the research of the GP approach to DA (Section 4.3.2) found emerging evidence that it has the potential to differentiate between the language learning potential of typically developing bilingual children and bilingual children with LI, few studies have been conducted. More specifically, no study to date has reported the classification accuracy of the GP approach to DA in discriminating the language learning potential of bilingual children with and without LI.

A review of the published literature found that no study has adopted a scripted MLE approach in combination with a structured GP approach to assess bilingual children's modifiability and therefore language learning potential. A highly structured and scripted approach to DA is anticipated to increase the ease of administration and scoring. This can in turn increase the clinical acceptability of DA as a preferred assessment for assessing bilingual children. The possible advantages and benefits are highlighted below:

1. The provision of MLE teaching instructions creates a naturalistic and engaging environment for testing to take place while reducing any test anxiety that children may have.

2. A scripted MLE approach allows for consistency of the procedure for all children. This also allows SLPs to provide mediation components within the DA without themselves having to undergo any additional training.
3. The provision of MLE instructions relies heavily on language as the mode of delivery (Camilleri & Bottling, 2013). A scripted MLE approach will allow the careful selection of simplified and developmentally appropriate teaching instructions for bilingual children who have at least six months of exposure to the language of instruction in school settings.
4. Using a GP framework allows examiners to score children's modifiability objectively rather than subjectively. This allows the interpretation of children's language learning potential to be impartial and reliable.

In summary, to increase the awareness and adoption of DA as a preferred alternative assessment approach for assessing the language learning potential of bilingual children in predominantly bilingual countries, there is a need to conduct local studies to provide evidence that DA can accurately differentiate the language learning potential of bilingual children with and without LI. At the same time, to increase the clinical acceptability of DA, it is helpful to explore the feasibility of a structured DA process that is relatively easy to administer and score. A review and critique of the literature on both MLE and GP approaches to DA suggest that a combination of both approaches to create a structured DA process may be a possible solution.

In the next section, the framework to gather validity evidence to demonstrate that the DA process developed in Phase Two of the study can evaluate and differentiate the language learning potential of bilingual children with and without LI will be discussed.

4.5 Validation Research in Language Testing

To increase the clinical acceptability and adoption of DA, test-users (i.e. SLPs, clinicians, professionals) need to be provided with clear evidence that the DA process developed can evaluate and differentiate the language learning potential of bilingual children with and without LI. In Phase Two, the DA process put together for the purpose of the study undergoes an evaluative process, known as validation, to provide evidence that DA as an alternative assessment approach can more accurately evaluate the language learning potential of bilingual children with variable language experience and therefore identify those at risk of LI.

The literature on language-testing shows that validation methods and frameworks have been developed to guide this process of evaluation (Chapelle & Voss, 2014). Before presenting the validity framework that is selected in this study to gather evidence to advocate the use of DA in bilingual clinical settings, it is useful to have an introductory understanding of the history of test validation.

The concept of validity has evolved over the last 50 years. Traditionally, validity is viewed as a property or characteristic of a test and is defined as the degree to which the test measures what it is designed to measure (Akbari, 2012). This general conceptualisation of validity was further broken down into different types of validity such as predictive validity, content validity, criterion-related validity, and construct validity (Cronbach & Meehl, 1955). These traditional views of validity focus mainly on the psychometric properties and the gathering of statistical evidence on the above-mentioned validity concepts to justify the quality of the test (Akbari, 2012; Chapelle & Voss, 2014). Although these concepts are still relevant and part of any sound test design process, they mostly ignore the interpretations of testing outcomes, social dimensions and consequences of test use.

Validity is currently viewed as the justification of the interpretation and uses of testing outcomes and is no longer just about the psychometric or statistical properties of a test (Akbari, 2012; Chapelle & Voss, 2014). In 1989, Messick proposed the concept of validity as a unified framework that highlights the social applications of a test. In other words, validity becomes an attempt to know what a test score means and its consequences rather than on focusing on what a test purports to measure (Akbari, 2012).

The core of Messick's conceptualization of validity is construct validity which he further deconstructs into six aspects to provide the framework for establishing the validity of a test: content, substantive, structural, generalizability, external and consequential (Akbari, 2012). Briefly, the *content* aspect evaluates the adequacy and relevance of the test's items and its content; the *substantive* aspect evaluates the theoretical rationale behind test takers' performance on the test item and that the test takers' performance is consistent with the theory. The *structural* aspect evaluates the fidelity of the scoring rubric of the test and whether it is in accordance with the theory behind the test. The *generalizability* aspect evaluates the extent to which the test scores can be generalised to and across samples, population groups and contexts, while the *external* aspect evaluates the extent to which a test correlates with other measures/test of the same skill or ability. Lastly, the *consequential* aspect evaluates the implications of the test score interpretations as well as the actual potential and consequences of test use (Akbari, 2012; Messick, 1995)

Although Messick's (1989, 1995) unified framework has been widely accepted, the main drawback is that the type and amount of evidence to be gathered for each aspect is often complex and unguided (Cook, Brydges, Ginsburg & Hatala, 2015; Davies & Elder, 2005). Therefore, test-developers and test-users are often unsure which source of evidence to prioritise and how to gauge progress, making validation a daunting task (Bachman, 2005; Chapelle & Voss, 2014; Cook, Brydges, Ginsburg, Hatala, 2015; Kane, 2012).

Davies and Elder (2005) have highlighted that if the notion of validity is to be credible and accessible to all, there is a need to simplify the processes of test validation, especially for low stakes testing. For example, the validation practices involved for high stakes tests and low stakes tests will be different. In high stakes testing (i.e. Tests of English as a Foreign Language) where test scores are used to make momentous decisions such as admission to university programs or licenses to practice a profession, there is a need for more rigorous validation practices that involves multiple inputs and professionals (i.e. professionals who have expertise in conducting validation research). Whereas, in low stakes testing (i.e. language screening assessment) where test scores are used to make screening, progress or achievement decisions, the validation process should be made less complex and achievable (Chapelle & Voss, 2014; Cook et al., 2015).

In the late 2000s, Kane (2006; 2012; 2013) proposed a step by step framework known as the Validity Argument Framework in gathering validity evidence when developing or/and appraising a test for its intended use. The Validity Argument Framework (Kane, 2006; 2013) guides the process and criteria for validation by 1) specifying the intended interpretations and uses of the test and; 2) highlighting the type of evidence to be gathered and evaluated in each of the key phases (i.e. inferences) of the framework. The advantage of the argument-based framework is that it is adaptable and can be applied to both high-stakes and low-stakes testing and across a wide range of qualitative and quantitative tests (Cook et al., 2015). This makes validation research seemingly straight-forward and achievable in the field of language and educational testing, especially with the development of low-stake tests (i.e. language screeners).

Before elaborating on Kane's Validity Argument Framework (2006; 2013) in Section 4.7, a summary of the type of validation research that has been conducted on DA in the field of language testing is first presented (Section 4.6).

4.6 Validating Dynamic Assessment: The Research Gaps

In some studies, researchers had attempted to provide validity evidence of their proposed DA tool/process. However, the focus of their validation efforts was mainly on the assessment itself in which the DA as a tool was validated with validity coefficients (i.e. the strength of the relationship between DA results and other criterion variables). The validity evidence was evaluated and established based on the psychometric properties/characteristics of the DA and not on the interpretations and uses of testing outcomes. For instance, Bain and Olswang (1995) investigated the validity of the DA they developed for evaluating young children's ability to produce two-word utterances. The construct validity of the DA was evaluated and established by reviewing the hierarchy of prompts provided in the DA. Results showed that as more support (more prompts) were provided through the cueing hierarchy, the participants were able to produce two-word utterances more accurately. Results also established the predictive validity for the DA where results showed that children who required less supportive prompts to produce two-word utterance in the DA demonstrated greater language change post-DA.

In a more recent study, Camilleri and Botting (2013) explored the validity of a DA of word learning ability developed in evaluating young children's lexical abilities. Concurrent validity of the DA was evaluated by correlating the children's performance on the DA with a static assessment of vocabulary. As both the DA of word learning ability and static assessment of vocabulary measured lexical abilities in young children, positive correlation coefficients were obtained, establishing the concurrent validity of the DA. Predictive validity was evaluated by correlating the children's performance on the DA with progress ratings of these children in classes and speech therapy provided by each child's key worker six months later. Positive correlations were found, establishing the predictive validity of the DA. In both studies, examples of validation efforts were focused on the property of the DA itself which were in line with Cronbach and Meehl's (1955) definition of validity.

Carlson and Heinz Wield (2000) have proposed that for DA the focus on validation efforts should be on its applicability and not on the psychometric properties of the DA itself. In other words, the justification of the interpretations to be made from the DA and the uses of the testing outcomes (i.e. test scores). However, to date, to the researcher's knowledge, no study has investigated or adopted an argument approach to validation to determine the validity of the interpretation and use of testing outcomes of a DA in evaluating the language learning potential of bilingual children with and without LI and therefore identifying those at risk of LI.

To contribute to the field of language testing for bilingual children, the Validity Argument Framework (Kane, 2006; 2013) was adopted to gather the evidence to support the proposed use and interpretation of the DA that was developed in Phase Two of the study. The purpose was to provide favourable evidence to increase the clinical acceptability of DA as a preferred assessment approach when assessing the language skills of bilingual children. An elaboration of Kane's Validity Argument Framework (2006; 2013) is provided in the next section (Section 4.7) with reference to the DA process that was developed in Phase Two of the study.

4.7 Validity Argument Framework

Kane's Validity Argument Framework (2006; 2013) provides a systematic approach to gathering and evaluating evidence at multiple points during the development of an assessment. The framework helps test-developers and test-users to prioritise the collection and evaluation of validity evidence (Cook et al., 2015). For Phase Two of the study, the validation process can assist in gathering evidence to support the interpretation and use of the proposed DA process as a potential assessment approach for evaluating the language learning potential of bilingual children and therefore identifying those at risk of LI.

Kane's (2012) argument-based framework is straight-forward and consist of two steps. Firstly, the Interpretation/Use Argument (IUA) is to be stated, followed by the validity argument in the

second step. In the first step, the IUA specifies the proposed interpretations and uses of assessment results by proposing a chain of inferences leading from observed performance to the conclusions and decisions based on the assessment scores (Kane, 2012). In the second step, the validity argument provides an evaluation (i.e. evidence) of the IUA's coherence and plausibility of the claims and assumptions under each inference (Kane, 2012).

To support the proposed interpretations and uses of assessment results, there is a need to identify the chain of inferences upon which the assessment results are to be based. Each inference in the IUA extends the interpretation or adds a decision regarding the use of the assessment (Kane, 2012). It starts from *Scoring* of a single observation to using the observed score to represent test performance that is consistent in the test setting (*Generalisation*), to drawing an inference as to how the test performance translates the test-taker's actual ability in the 'real-world' (*Extrapolation*) (Cook et al., 2015; Kane, 2006; 2013).

Kane identified two type statements (i.e. claims and assumptions) as the framing statements under each inference (Chapelle, Enright & Jamieson, 2010). Claim statements are generally held principles, the rule of thumb or established procedure. The claim rests on assumption statements that need to be supported in order for the inference to be made. Assumptions can guide the collection of relevant evidence to support the interpretation and use of the assessment.

Kane intended the Validity Argument Framework (2006; 2013) to be a critical validation process to objectively gather and evaluate the evidence for the proposed interpretations and uses of a developed assessment. However, the framework can also be used to gather development evidence when exploring or creating a new assessment process/tool. At the development stage, the purpose of gathering evidence is to show that the interpretation and use of the assessment process/tool that is being explored are justified. Therefore, the evidence

presented will tend to have a confirmationist bias to advocate for the proposed use and interpretation of the assessment process/tool (Kane, 2006; 2013).

When an assessment process/tool is being developed, the materials and procedures of the assessment are expected to be developed in a way that supports the proposed interpretation and use of the assessment. Although Kane did not explicitly define the construct of the assessment, it is pre-assumed that the test-developers would have selected materials and procedures as a natural part of the process to make a case for the validity of the proposed interpretations and uses of the assessment (Kane, 2013). In other words, the intended interpretation of the assessment results is based on a domain description inference that can be used to infer a decision (i.e. typically developing or LI). However, Chapelle and colleagues (2010, 2014) reasoned that the domain description is critical to evaluate interpretation and use of an assessment. Therefore, they have added an inference for domain description (*Domain Definition)* to the Validity Argument Framework (Kane, 2006; 2013) in their studies.

In Phase Two of the study, the DA process is to be developed to explore whether DA is a more accurate assessment approach compared to a standardised assessment of vocabulary in evaluating the language learning potential of bilingual children with and without LI and identifying those at risk of LI. The validation process will begin by stating the proposed interpretations and use of the DA which is - when a child is observed to score poorly on the DA, the child is likely to have poor language learning potential and is at risk of LI (Figure 4-2).

Next, the IUA of the DA is to be stated by listing the key claims and assumptions under each inference (Figure 4-2 and Table 4-1 below). Evidence then has to be gathered to support the claims and assumptions under each inference. As mentioned previously, during the development phase, the evidence gathered will be of confirmationist bias as the purpose of

gathering validity evidence is to advocate the use of DA in evaluating the language learning potential of bilingual children and therefore identifying those at risk of LI. If the evidence gathered is favourable, the claims and assumptions under the specific inference will be accepted, and the next group of evidence to be evaluated for the next inference would proceed. However, if the evidence is unfavourable, the DA process would be revised (Figure 4-3).

Figure 4-2. Interpretation and use of the DA.

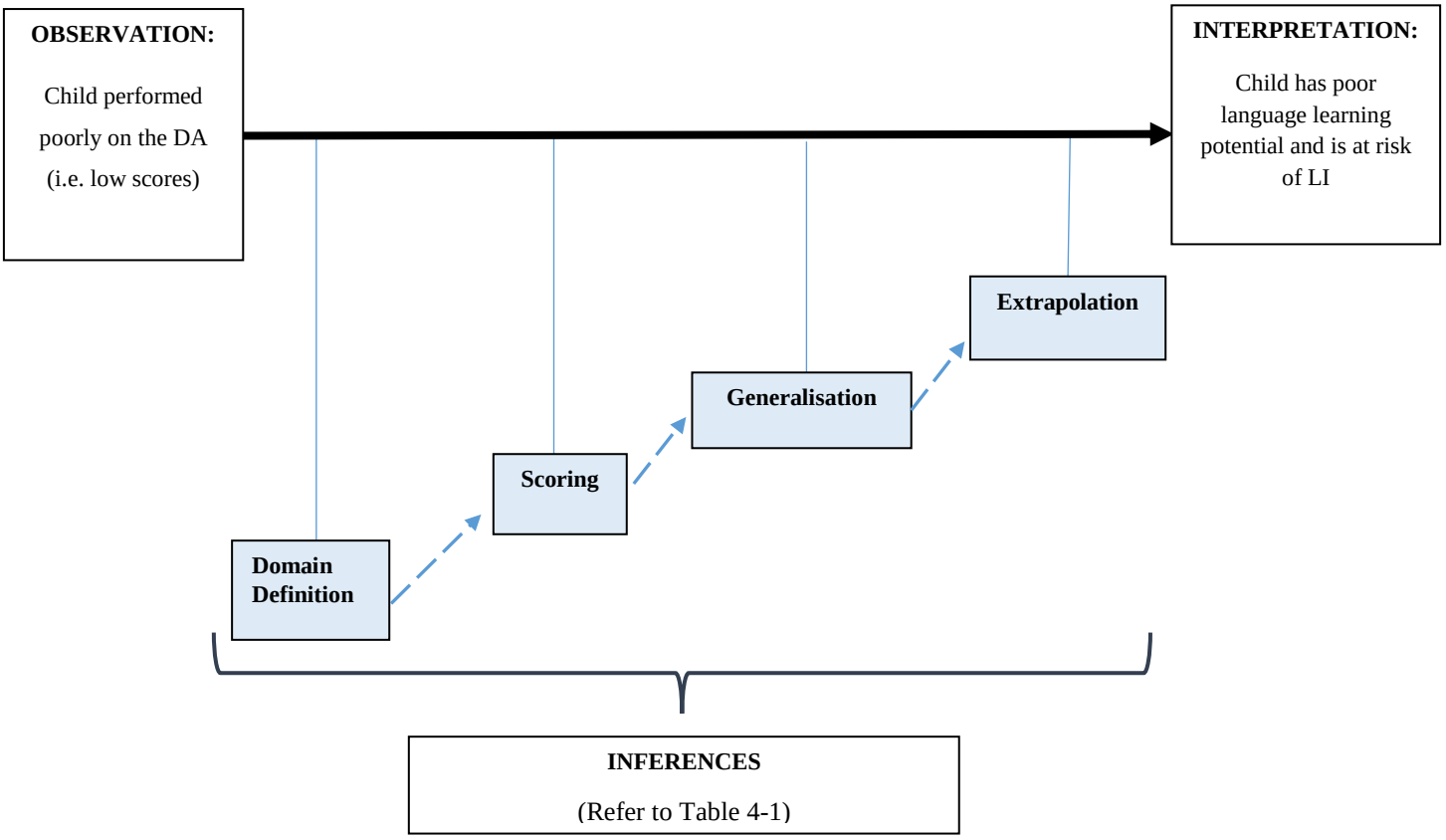
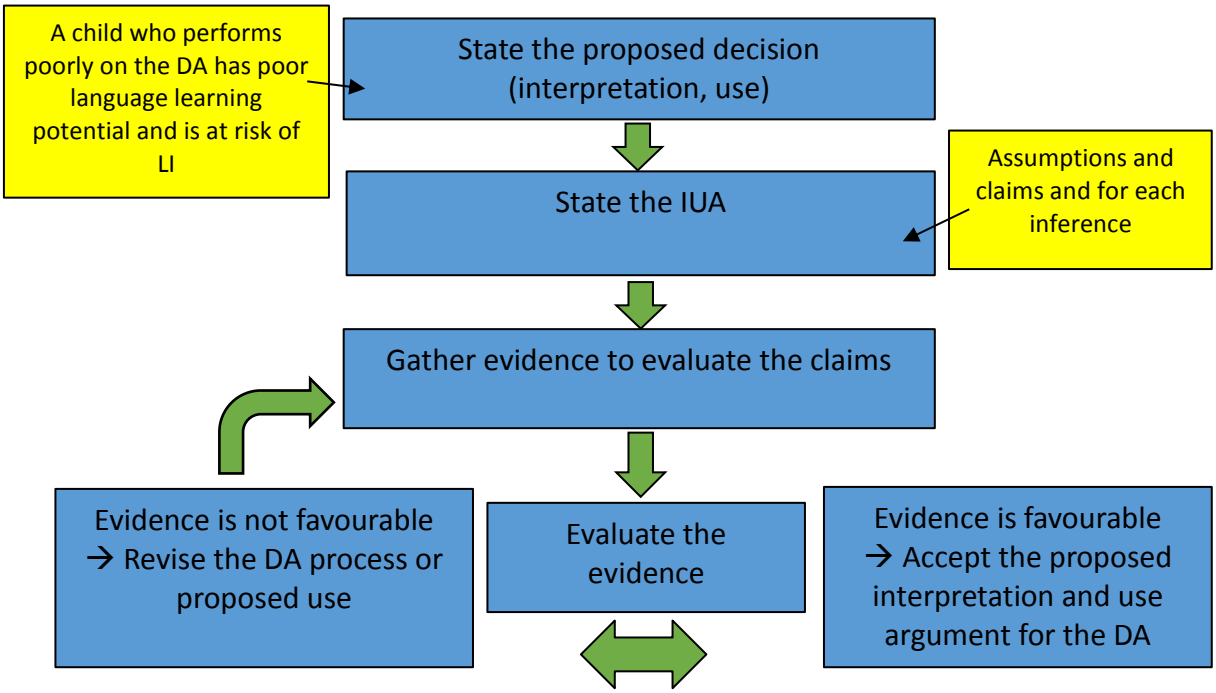


Figure 4-3. Validation process for the DA.



The interpretive argument for the DA contained the following claims: 1) Observations of performance on the DA reveal a bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI; 2) Observations of performance on the DA are evaluated to provide observed scores reflective of language learning potential; 3) Observed scores on the DA reflect language learning potential that are consistent among bilingual children with the same language developmental profiles regardless of their prior language experience; 4) The construct of language learning potential as assessed on the DA predicts scores that correspond to bilingual children's language developmental profiles.

A list of claims and assumptions under each inference to be evaluated in the process of developing the DA for the current study are presented in Table 4.1.

Table 4-1. Claims and assumptions to be evaluated for the proposed DA.

Inference	**Claims**	**Assumptions**
Domain Definition	Observations of performance on the DA can reveal bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI.	The language task selected to be evaluated in the DA is developmentally appropriate and least biased for evaluating language learning potential among bilingual children.
Scoring	Observations of performance on the DA task are evaluated to provide observed scores reflective of language learning potential.	The task procedure provides opportunities for language learning to take place. The task administrative conditions are appropriate for observing and scoring language learning opportunities. The rubric for scoring responses are appropriate for providing evidence of language learning ability.
Generalisation	Observed scores on the DA are consistent among bilingual children with the same language developmental profiles regardless of their prior language experience.	The scores obtained on the DA between bilingual children with and without LI are differentiated. The scores obtained on the DA by bilingual children with the same language developmental profiles are consistent regardless of their prior language experience.
Extrapolation	The construct of language learning potential as assessed on the DA predicts scores that correspond to bilingual children's language developmental profiles.	The scores obtained on the DA can accurately predict bilingual children's language developmental profiles.

The evidence to support the claims under each inference of the Validity Argument Framework (Kane, 2006; 2013) will be discussed in the subsequent chapters of the dissertation.

4.8 Chapter Summary

There is emerging evidence that DA is a more appropriate assessment than static assessments in evaluating the language ability of bilingual children to identify those at risk of LI. However, the research conducted to date has been with bilingual children from minority communities in predominantly monolingual countries. The feasibility and utility of DA in identifying bilingual children at risk of LI in predominantly bilingual countries have not been investigated.

MLE and GP approaches to DA have been found individually to be appropriate approaches for evaluating children's modifiability and distinguishing LI from typical development. However, there are limitations in each approach that have impacted on the clinical acceptability of DA. A structured approach to DA that combines the elements of both approaches (i.e. scripted MLE, GP framework in evaluating modifiability) has the potential for DA to be more readily adapted and used in bilingual clinical settings.

Last not but least, for test-users (i.e. SLPs) to be confident in using DA to identify bilingual children at risk of LI, validity evidence to justify the interpretation and use of the DA has to be collected and evaluated. The Validity Argument Framework by Kane (2006; 2013) provides an outline for the chain of validity evidence to be collected and evaluated to support the proposed interpretation and use of the DA.

In the next two chapters (Chapter 5 and Chapter 6), the development of the DA that occurred in Phase Two of the study will be elaborated on. Before the development of the DA began, the first inference - *Domain Definition* under the Validity Argument Framework (Kane, 2003;2006) had to be supported to make the first claim - *Observations of performance on the DA reveal a bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI*. In the next chapter, the evidence to support the suggestion that

the language task selected to be evaluated in the DA is relevant for evaluating language learning potential and the identification of LI among bilingual preschool children will be presented.

Chapter 5 Phase Two: Literature Review Part B

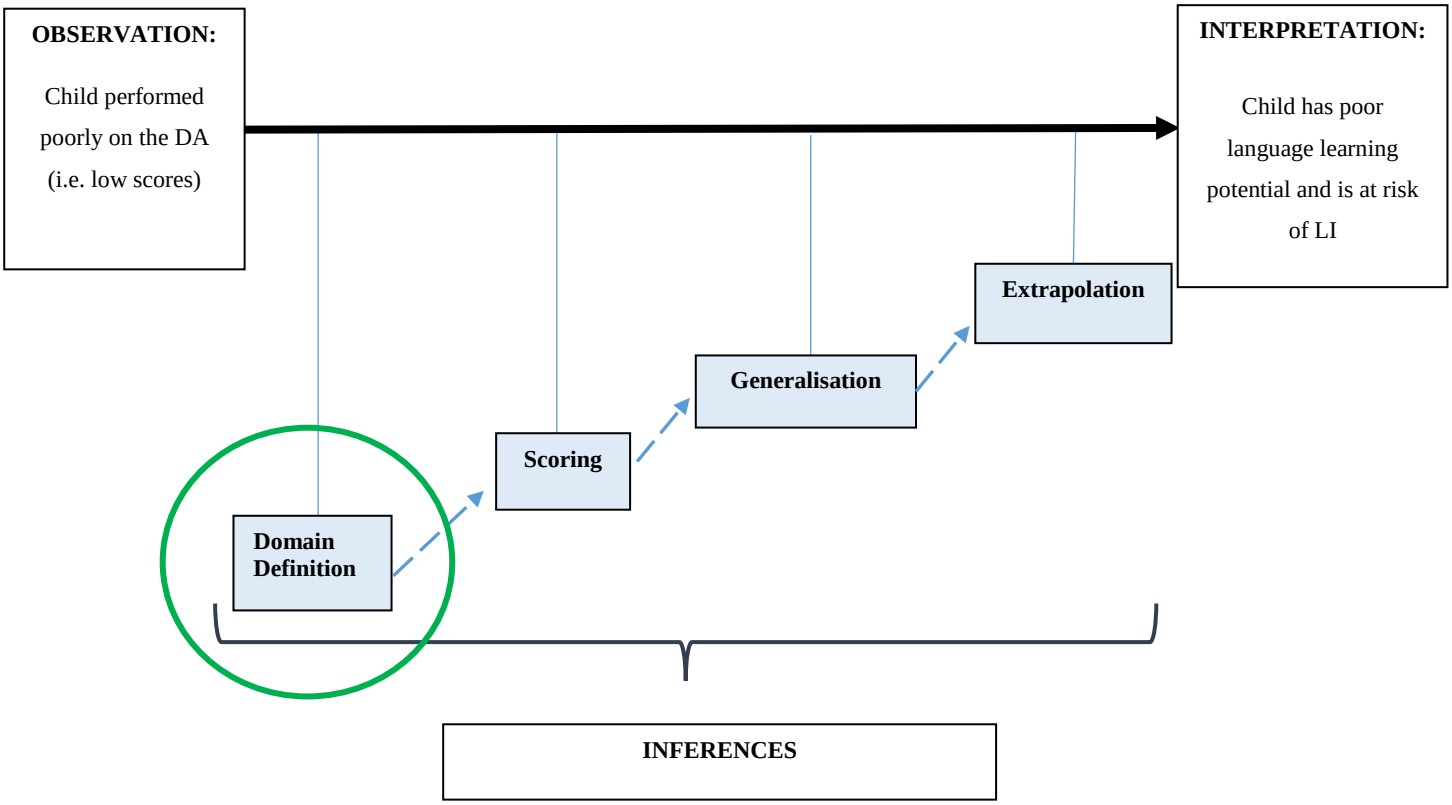
5.1 Introduction

The primary aim of Phase Two of the study was to develop and explore a DA process that could assist SLPs working with predominantly bilingual populations in evaluating the language learning potential of bilingual preschool children. More specifically, whether the DA process could validly identify bilingual preschool children at risk of LI as compared to a commercially-available standardised assessment. The DA process was designed to evaluate the language learning potential of preschool children as it is vital to identify LI at a young age for timely and responsive intervention to take place to improve language and related outcomes (Prelock, Hutchins & Glascoe, 2008).

In the preceding chapter, the Validity Argument Framework (Kane 2006, 2013) was identified as the framework for gathering evidence to support the proposed interpretation and use of the DA process that was developed and evaluated in Phase Two (Figure 5-1). Prior to developing the DA process, an appropriate language task had to be first identified. The assumption was that the language task evaluated in the DA had to be one where poor performance (i.e. a low score) was a behavioural indicator of LI.

In this chapter, a review of the relevant literature is presented to provide evidence to support the first claim (*Observations of performance on the DA can reveal bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI*) under the *Domain Definition* of the Validity Argument Framework (circled in green in Figure 5.1). First, an overview of the theoretical accounts of LI is provided.

Figure 5-1. The Validity Argument Framework.



5.2 Theoretical Accounts of Language Impairment

The theoretical accounts of LI were reviewed to identify a suitable language task to be assessed in the DA. The language task selected needed to be one where poor performance is a behavioural indicator of LI.

Current theories of LI can be classified into two broad categories: linguistically-based accounts and processing-based accounts. This distinction was first introduced by Chomsky (1957; cited in Elli Weismer & Evans, 2002) who conceptualized the differentiation between a person's knowledge of grammatical rules (*linguistic*) and the language processing functioning of a person which is influenced by non-linguistic factors such as memory and attention (*processing*).

5.2.1 Linguistically-based (deficit) Accounts

In linguistically-based accounts, it is proposed that children with LI present with overall delays in language development with the main language deficit reflected in the children's underlying

grammar (Bedore & Peña, 2008; Paradis, 2007; 2010). Children with LI are described to have marked difficulties acquiring/marking verb tense morphology (e.g. copula forms, auxiliary forms, verb inflections that mark agreement and/or tense: Conti-Ramsden & Windfuhr, 2002; Leonard, Eyer, Bedore, Grela, 1997; Oetting & Horohov, 1997; Rice & Wexler, 1996). There are a number of theories that differ with respect to which features of the children's grammar are impaired (i.e. Agreement-deficit Account: Clahsen, 1989; Limited Grammatical Rule-learning Capacity: Gopnik & Crago, 1991; Extended Optional Infinitive: Rice & Wexler, 1996).

For instance, one of the most studied and documented hypotheses is the Extended Optional Infinitive (EOI) profile by Rice and Wexler (1996) in which children with LI are described to have difficulties acquiring and marking finite tense morphology. It is termed as *Extended* Optional Infinitive as children with LI are observed to demonstrate Optional Infinitive (OI) profiles that are only seen in younger typically developing children. Typically developing children are expected to acquire and master finite verb tense marking by the age of four (i.e. English language: Rice, Wexler & Cleave, 1995; Rice, Wexler & Hershberger, 1998; French language: Paradis & Crago, 2001). However, children with LI continue to display OI profiles even after the age of seven and may never fully master the marking of finite verb tense (Rice et al., 1995; 1998). The EOI profile is interpreted within the context of Rice's (2004) maturational model of language growth. In the model, language growth is initiated by genetically timed mechanisms. For typically developing children, the mechanisms occur as programmed to support language growth. For children with LI, the presence of language delay results in disruptions to the genetically timed mechanisms within the linguistic system. The EOI profile occurs when there is a late onset of verb tense marking, or/and when there are disrupted growth patterns between finiteness marking morphology and other grammatical morphology.

The difficulty in marking verb tense is widely proposed as a clinical behavioural marker for identifying LI among monolingual children across languages. Therefore, grammatical-type tasks are often used to evaluate the children's language skills where poor performance is used as a behavioural indicator for LI. For instance, in the English language, studies (e.g. Rice & Wexler, 1996; Bedore & Leonard, 1998; Oetting & Horohov, 1997) have found that monolingual English-speaking children with LI between the ages of four to eight demonstrated poor accuracy and non-mastery of verb tense marking compared to their same-aged typically developing peers who demonstrated high accuracy and mastery of verb tense marking. Similarly, the difficulty in marking verb tense was also found in monolingual French-speaking children with LI (Paradis & Crago, 2001), monolingual German-speaking children with LI (Rice, Noll & Grimm, 1997), monolingual Italian-speaking children (Bortolini, Caselli & Leonard, 1997) and monolingual Spanish-speaking children with LI (Bedore & Leonard, 2001).

However, studies have also shown that there are similarities in the grammatical error patterns made between typically developing bilingual children and their monolingual counterparts with LI (Brebner et al., 2016; Håkansson, 2001; Paradis, 2005; Paradis & Crago, 2000; Paradis, Rice, Crago & Marquis, 2008; Thordardottir, 2015). For instance, Paradis (2005) investigated the performance of 24 typically developing bilingual preschool children who were learning English as their second language on their production accuracy and errors on a series of English grammatical morphemes. Results showed that these bilingual children, who had been learning the English language for less than two years, presented with accuracy rates and grammatical errors patterns that were similar to those reported for same-aged monolingual English-speaking children with LI. Paradis and colleagues (2008) then compared the same group of bilingual children to monolingual English-speaking children with LI and typically developing monolingual English-speaking children and found that there were similarities in tense acquisition and grammatical errors patterns across all three groups. The results from both

studies suggest that in practice, it is almost impossible to determine whether the grammatical errors observed in young bilingual children who are still learning the language are indicative of LI. In other words, the difficulty in marking verb tense may not be a reliable behavioural indicator for LI among bilingual children.

The same observations were also reported in languages other than English. Studies found that typically developing bilingual children learning French or Swedish had strikingly similar grammatical error patterns in French or Swedish when compared to their monolingual counterparts with LI respectively (Paradis & Crago, 2000; Håkansson, 2001). In other words, if the deficit in marking verb tense is used as a behavioural indicator for LI across languages, there will be an over-identification of bilingual children with LI when they, in fact, have normal language ability.

Furthermore, grammatical development in each language in bilingual children is strongly influenced by the amount of input/exposure in each language. Thordardottir (2015) examined the effect of bilingual input in each language on the grammatical development of typically developing English-French bilingual children in each language. She found that the children who had less exposure to the English language (<40.0%) in comparison with the French language displayed grammatical error patterns in English that were consistently reported for monolingual English-speaking children with LI. Thordardottir also concluded that difficulty in marking verb tense in the English language is not an appropriate indicator of LI in bilingual children who had less exposure and are still learning the language.

In another study, Brebner and colleagues (2016) investigated the rate and pattern of verb tense marking in the English language among typically developing Singaporean English-Mandarin bilingual preschool children who were either reported to have had more exposure to the English language (i.e. English dominant) or more exposure to the Mandarin language (i.e. Mandarin

dominant). The results found that the group of English dominant children demonstrated faster and more consistent marking of verb tense compared to the group of Mandarin dominant children who demonstrated almost no evidence of verb tense marking even at kindergarten. Most importantly, the results highlighted that even for the group of English dominant bilingual children, they demonstrated acquisition of verb tense marking at a much later rate compared to what is expected of monolingual speakers of the English language. This was likely due to the influence of the Mandarin language, a tonal language that possesses very little inflectional or derivational morphology (Fung, 2008). This suggests that even young bilingual children who were predominantly exposed to the English language acquired verb tense marking at a different rate compared to monolingual norms. Therefore, the use of absence or delayed marking of verb tense marking is not an accurate indicator of LI for young bilingual children.

A review of the studies in this section has summarised that the acquisition and mastery of verb tense marking among bilingual children occurs at a different rate and developmental trajectory compared to monolingual developmental norms. Similarly, in a review of studies examining language development in bilingual children, Paradis (2016) concluded that bilingual children often take more than three years to approach age-expected monolingual abilities for most linguistic subdomains (i.e. morphology, vocabulary) in each language they speak. While learning the language, bilingual children may display grammatical error patterns that are observed in monolingual children with LI. Therefore, the absence or delay of marking verb tense as a behavioural indicator for LI under the linguistically-based account cannot be reliably used to identify LI among young bilingual children. In other words, grammatical-type tasks are likely not appropriate to be evaluated in the DA context for the identification of LI among young bilingual children.

5.2.2 Processing-based (deficit) Accounts

The limitations in using linguistically-based behavioural indicators (i.e. poor verb tense marking) to identify LI in bilingual children have turned the focus towards processing-based behavioural indicators instead. Under the processing-based account, it is assumed that language difficulties observed (i.e. grammatical deficits, limited vocabulary) are secondary to limitations or deficits in cognitive processing capacity that are used for processing linguistic information as well as other cognitive operations (Bedore & Peña, 2008; Paradis, 2007; 2010). These hypotheses can be grouped into those that assume that deficits are in specific processes (e.g. *Deficits in phonological working memory*: Baddeley, 2003; Gathercole & Baddeley, 1990, *Surface Hypothesis*: Leonard, 1998) and those that assume deficits in general information processing functioning (e.g. *Restrictions in working memory capacity*: Elli Weismer, Evans & Hesketh, 1999; Montgomery, 2002, *Slower speed of processing*: Miller, Kail, Leonard & Tomblin, 2001). For example, Baddeley and colleagues (1990; 1998; 2003) proposed that deficits in the phonological loop of the working memory underlie the cause of LI. The phonological loop helps to encode and rehearse new auditory phonological information before it becomes embedded in the long-term memory. It is proposed that children with LI have deficits in auditory perception that result in them not being able to encode a sufficient amount of novel phonological information (i.e. new word) and/or limitations in capacity that only allow them to encode a limited amount of novel phonological information (Alt, 2011). The limitation in the phonological loop can result in incorrectly formed phonological information (i.e. incorrectly learned words), partially formed phonological information (i.e. partially learned words), or phonological information that is not encoded/learned at all. When compared to typically developing children who have intact cognitive processes, children with LI will have less well-formed representations of the language that impact on comprehension and production as well as further language learning (Alt, 2011).

Non-linguistic processing tasks such as perceptual processing, fine motor sequencing, cognitive imaging tasks have accounted for difficulties observed in monolingual children with LI when compared to typically developing monolingual children (i.e. Kohnert & Windsor, 2004). However, besides the study by Kohnert and colleagues (2009) that found that non-linguistic processing tasks (mental rotation, auditory pattern matching, choice visual detection and visual form completion) differentiated the performance of Spanish-English bilingual children and monolingual children with LI. No study has studied whether the same tasks can differentiate the non-linguistic processing performance of bilingual children with and without LI.

Similarly, the limitations in cognitive processing capacity have also accounted for difficulties observed in monolingual children with LI in language processing tasks where new language information is to be learnt and recalled. For example, limitations in short-term phonological memory result in difficulties in discriminating, encoding and producing non-words in non-word repetition tasks (i.e. Dollaghan & Campbell, 1998; Ellis Weismer & Evans, 2002; Estes, Evans, Else-Quest, 2007). Similarly, limitations in working memory result in poor sentence recall/repetition in sentence repetition tasks (i.e. Archibald & Joanisse, 2009; Conti-Ramsden, Botting & Faragher, 2001; Thordardottir et al., 2011). Last but not least, limitations in working memory also result in poor word learning performance in novel word learning tasks (i.e. Kan & Windsor, 2010; Nash & Donnaldson, 2005). In particular, findings from the literature have identified poor performances in non-word repetition tasks (e.g. Bishop, Noth & Donlan, 1996) and sentence repetition tasks (e.g. Conti-Ramsden et al., 2001) as potential behavioural indicators for LI for monolingual children.

There is a growing body of evidence suggesting that non-word repetition tasks and sentence repetition tasks can differentiate between the performance of typically developing monolingual children and monolingual children with LI. At the same time, there is emerging evidence that

the same tasks can be used to differentiate the language processing performance of bilingual children with and without LI. However, the evidence from the literature is still mixed. To illustrate this point, studies and their findings that evaluate the performance of bilingual children with and/or without LI on non-word repetition and sentence repetition tasks are presented in the next few paragraphs.

Kohnert and colleagues (2006) examined the performance of three groups of school-aged children: monolingual English speaking children with LI, typically developing monolingual English speaking children and typically developing English-Spanish bilingual children on a non-word repetition task. The non-words were two to six syllable non-words that followed English phonological rules and performance scores were tabulated as percentages of phonemes produced correctly. Results showed that although the group of typically developing bilingual children on average produced a statistically significantly higher percentage of correct phonemes (87%) compared to the group of monolingual English speaking children with LI (81%), they still produced a statistically significant lower percentage of correct phonemes compared to the group of typically developing monolingual English speaking children (92%). When the range of scores obtained by each group were further analysed, the group of bilingual children were found to obtain a range of scores (15-93%) that overlapped the scores obtained by the group of monolingual English speaking children with LI (23-88%) and typically developing monolingual English speaking children (30-100%). The authors calculated the likelihood ratios for four levels of cut-off scores (<72%, 73-81%, 82-92%, >93%) with the maximum diagnostic change to determine classification accuracy. They concluded that poor performance in non-word repetition tasks may have appropriate high specificity that can rule out LI among bilingual children when a high cut-off score (>93%) that was above the average obtained by typically developing English-speaking monolingual children (92%) was used. However, because of the broad range of scores obtained by the group of typically developing

bilingual children, the lowest cut-off score (72%) was not sensitive enough to rule in or identify LI among bilingual children.

Performance on non-word repetition tasks is also not completely independent of linguistic bias (De Lamo White & Jin, 2011). Gutiérrez-Clellen & Simon-Cereijido (2010) examined the performance of 95 typically developing Spanish-English bilingual preschool children and 49 Spanish-English preschool children with LI on non-word repetitions in both languages (i.e. 16 non-words in English, 20 non-words in Spanish). Although results showed that the group of bilingual children with LI demonstrated a significant deficit in non-word repetition tasks compared to their typically developing peers, a considerable number of typically developing bilingual children did not obtain passing scores, especially in their non-dominant language. Likewise, Summers and colleagues (2010) explored the performance of 62 typically developing Spanish-English bilingual children on non-word repetition tasks in both languages and concluded that their performance in non-word repetition in each language was significantly correlated to the language experience they each have in the language. The findings from both studies demonstrated that: 1) the performance of bilingual children in non-word repetition tasks is influenced by their prior language experience, and 2) the clinical accuracy of non-word repetition tasks to differentiate between bilingual children with LI and typically developing bilingual children is likely dependable on the language the task was presented in (i.e. whether it is presented in the language that the bilingual child has more exposure or less exposure to).

Similarly, Thordardottir and Brandeker (2012) examined the performance of 84 typically developing French-English bilingual preschool children, differing in the amount of exposure in both languages, on non-word repetition and sentence imitation tasks in each language. The results found that except for the non-word repetition tasks in French, the performance in the other three tasks (non-word repetition in English, sentence-imitation in French, sentence imitation in English) was highly associated with the amount of exposure they had received in

each language. The researchers suggested that because sentence imitation tasks use real words and sentences while non-word tasks use non-words that still follow the phonotactic rules of each language, the influence of prior language experience was expected. In the second phase of the same study, the same non-word repetition tasks and sentence imitation tasks in French were administered to monolingual French-speaking and bilingual French-English children with and without LI (14 children in four groups). Although both the non-word repetition and sentence imitation tasks differentiated children with LI from children with typical development regardless of their language profiles (i.e. monolingual or bilingual), the classification accuracy for both tasks for French-English bilingual children (Non-word repetition: Sensitivity: 85%, Specificity: 79%; Sentence imitation: Sensitivity: 92%, Specificity: 57%) was still lower compared to that of the monolingual French speaking children (Non-word repetition: Sensitivity: 92%, Specificity: 100%; Sentence imitation: Sensitivity: 93%, Specificity: 86%). Furthermore, while the sensitivity levels for the non-word repetition and sentence imitation tasks to identify LI in French-English bilingual children were acceptable, the specificity levels to rule out LI in French-English bilingual children were not.

In summary, although language processing tasks may not be completely independent of linguistic bias, they are still a less biased option compared to linguistically-based language tasks. Moreover, the evidence for the diagnostic value of language processing tasks such as non-word repetition and sentence imitation to validly identify LI among bilingual children is also emerging. In a DA context, the use of language processing tasks offers an advantage over linguistically-based language tasks for identifying LI among bilingual children. This is because language processing tasks can evaluate bilingual children's ability to process/learn the new language information while minimising the influence of prior language experience. This is parallel to the purpose of DA which is to evaluate children's language learning potential.

There is potential in evaluating bilingual children's language learning potential in a DA context using a language processing task. To consider an aspect of processing limitation observed in children with LI and to extend it to a language processing task could assist in the identification of LI in young bilingual children, there is a need to take into account the early developmental stages of language acquisition in young children (Ellis Weismer & Evans, 2002). One of the first sublinguistic domains that young children acquire is vocabulary knowledge, and an appropriate language task that could be evaluated under the processing-based account that reflects this developmental stage is word learning ability. Bedore and Peña (2008) have previously proposed that as difficulties relating to word learning are consistent across languages in children with LI regardless of their language experience, it may serve as a potential behavioural indicator for identifying LI among bilingual children.

In the next section, the literature on word learning in children with and without LI is discussed. This is to determine whether word learning may be a suitable choice of language processing task to be evaluated in a DA context.

5.3 Language-processing Measure: Word Learning Task

Word learning occurs as a gradual process where the child hears a novel word and after a few or repeated exposures, processes both linguistic (i.e. phonological, semantic, syntactic information) and non-linguistic information (i.e. context in which the word is introduced) to create a mental representation of the word (Ellis Weismer & Evans, 2002; Kapantzoglou et al., 2012). Word learning occurs in two stages: fast mapping and slow mapping (Carey & Bartlett, 1978). Initially, after a single or few exposures to the novel word in a single context, the child creates only a partial mental representation of the word (fast mapping). Thereafter, subsequent exposures to the word in various contexts allow the child to strengthen the mental representation of the word (slow mapping).

Word learning is a process that relies heavily on cognitive processes (Kan & Windsor, 2010). Research has shown that limitations in processing capacity and/or processing skill (i.e. deficits in short-term phonological memory, difficulties in phonological encoding, deficits in retrieval, deficits in associating semantic information to form and vice versa) can account for the difficulties observed in children with LI in processing, storing and retrieving linguistic and non-linguistic information about the new words (Alt, Plante & Creusere, 2004, Ellis Weismer & Evans, 2002, Kan & Windsor, 2010). For instance, Dollagham (1987) investigated the word learning ability of 11 typically developing preschool children and 11 preschool children with LI on learning a single novel word after a single exposure. The study found that there was a significant difference in word production performance between both groups of children. Even in a supportive word learning context where phonological and semantic cues were provided, Gray (2005) found that preschool children with LI exhibited both restricted phonological and semantic processing abilities that led to a significantly lower number of novel words identified and produced when compared to the age-matched typically developing peers. In a more recent study, Jackson and colleagues (2016) investigated the performances of 5- to 6-year-old preschool children with and without LI on a novel word learning and non-word repetition task. Results revealed that children with LI obtained significantly lower word production scores in comparison with their typically developing counterparts. This pattern was also observed in the non-word repetition task where children with LI also obtained significantly lower scores compared to their typically developing counterparts, indicating limitations in short-term phonological memory capacity. The authors concluded that limitations in short-term phonological memory capacity lead to poor word learning ability and poor non-word repetition performance among children with LI. The different examples from novel word learning studies provide supportive evidence that limitations in processing capacity/skills result in poor word learning ability among children with LI.

The overall findings from novel word learning studies indicate that monolingual children with LI tend to comprehend and produce fewer words in comparison with their typically developing peers (Alt et al., 2004; Rice et al., 1992, 1994; Gray 2003, 2004; Jackson, Leitao & Claessen, 2016; Kan & Windsor, 2010; Oetting, Rice & Swank, 1995). For instance, Kan and Windsor (2010) compared the difference in novel word learning performance of children with LI and typically developing children across 28 studies and concluded that monolingual children with LI indeed demonstrated poor word learning skills compared to the same age monolingual typically developing peers. More specifically, on average, across 28 studies, children with LI demonstrated word learning skills that were at least half a standard deviation below that of their age-matched typically developing peers. The difference in word learning ability between preschool children with LI (aged 5;11 and below) and their typically developing age-matched peers was also found to be larger compared to studies that examined word learning ability in school-aged children with and without LI (aged 6;0 to 12;3). This is because, in younger children, cognitive processing skills are still developing which impact on their ability in processing linguistic information in context (Dixon, Salley & Clements, 2006). It is likely that the deficits in cognitive processing skills in younger children with LI further exacerbate their limitations in processing linguistic information leading them to demonstrate even poorer word learning performance across studies. For school age-children with LI, however, the maturation of the cognitive system, language and school experience, may assist them in developing compensatory strategies to redress their deficits in cognitive processing to support word learning (Kan & Windsor, 2010). In other words, poor word learning performance is likely an appropriate behavioural indicator for identifying LI among young preschool children but not so with older school-aged children.

The research in novel word learning ability among bilingual children with and without LI is still sparse. Nonetheless, there is emerging evidence that bilingual children with LI also exhibit

poorer word learning ability in comparison with their age-matched, typically developing bilingual peers. Kapantzoglou and colleagues (2012) examined the word learning ability of 4- to 5-year-old predominantly Spanish-dominant bilingual children with and without LI on a word learning task (3 novel words) using a DA paradigm. Overall, the means for word identification and production scores obtained by the group of LI bilingual children were all lower in comparison with the group of typically developing bilingual children. In other words, similar to the findings of the monolingual studies, bilingual children with LI made associations between phonological and semantic representations of new words more slowly than their typically developing bilingual peers.

In summary, monolingual children with LI consistently exhibit poorer word learning ability compared to their typically developing, age-matched peers. Although there has only been one study (Kapantzoglou et al., 2012) so far that examined the difference in novel word learning performance between bilingual preschool children with and without LI, results from that study alone provided some emerging evidence that poor performance in novel word learning tasks has the potential to be used as a behavioural indicator to identify LI in young bilingual preschool children. Therefore, word learning is a suitable task to be evaluated in a DA context to 1) differentiate the language learning potential between young bilingual children with and without LI, and 2) use as a behavioural indicator where observation of poor word learning ability indicates that the child is at risk of LI.

To support poor word learning performance as an appropriate language processing task to identify bilingual preschool children at risk of LI in a DA context, the factors that may influence and maximise the difference in word learning performance between children with and without LI must be further explored and understood. In the next section (Section 5.4), the factors that are known to influence word learning performance in children are reviewed and presented.

5.4 Factors that Influence Word Learning Ability in Children with and without Language Impairment

Although children with LI generally have poorer word learning ability when compared to typically developing children, there is a considerable variability in the extent of the group difference in word learning ability between children with and without LI across studies. Some studies have reported significant group differences across all areas of word learning (i.e. Gray, 2003: significant group differences in both word identification and production scores observed) while some studies reported no significant group difference in at least one area of word learning (i.e. Dollaghan, 1987: significant group difference in word production scores observed but not in word identification). This variability is likely due to the methodological variations in task characteristics of word learning tasks.

It is essential to consider the task characteristics of word learning tasks in order to maximise the difference in word learning performance between bilingual children with and without LI in a DA context. Kan and Windsor (2010), in their meta-analysis of 28 studies, have highlighted that task characteristics such as word type and the number of presentations in input have differential influences on the word learning performance of children with and with LI. The characteristics of the task should provide sufficient support for typically developing bilingual children to learn new words easily but remain challenging for bilingual children with LI to reveal their processing limitations. This is to ensure that both floor and ceiling effects are avoided. The task characteristics of particular interest to assess word learning ability in a DA context are word type and number of presentations.

5.4.1 Word Type

Most word learning studies have selected non-words or highly unfamiliar real words as stimuli to evaluate word learning ability in children. Similar to non-word repetition tasks and sentence imitation tasks, children's performance on word learning tasks can be influenced by previous

language exposure if common and/or familiar real words are used as target words (Kan & Windsor, 2010, Kapantzoglou et al., 2012). However, even with unfamiliar real words, it is impossible to tell if the child has had any previous exposure the words which may impact on their performance on word learning tasks. When non-words are used as stimuli instead, the influence of previous language exposure on word learning performance is minimised. Therefore any word learning behaviours can be attributed to the cognitive-linguistic processes that are involved in word learning and not to prior language experience.

Additionally, when non-words are used, phonotactic characteristics such as phonotactic probability and neighbourhood density can be controlled. Phonotactic probability refers to the likelihood of the phonemes and phoneme sequences of the word occurring in a given language (Storkel, 2001) whereas neighbourhood density refers to the number of real words in a given language when one phoneme is added to, deleted or changed in the word (Alt & Plante, 2006). Phonotactic probability primarily affects processing of sublexical representations (i.e. sound sequences) while neighbourhood density affects lexical representations (Vitevitch, Luce & Pisoni & Auer, 1999). Phonotactic probability and neighbourhood density were found to have different effects on word learning performance among children with and without LI (Gray, Pittman & Weinhold, 2014; McKean, Letts & Howard, 2013).

Studies have shown that non-words with high phonotactic probability facilitate word learning in typically developing preschool children (i.e. Storkel, 2001; 2003, Gray et al., 2014). However, for preschool children with LI or phonological impairment, no phonotactic probability effect has been found instead. For example, Gray and colleagues (2014) investigated the effects of phonotactic probability on word learning performance in preschool children with and without LI. Results showed that for both word identification and word production performance, typically developing children benefited from a high phonotactic probability effect whereas children with LI demonstrated no phonotactic probability effect.

Gary and colleagues (2014) proposed that because cognitive processing skills are intact in typically developing children, non-words with high phonotactic probability are likely held in the working memory longer and more accurately. This allows phonological processing to be facilitated which allows encoding and storage of novel words with high phonotactic probability to happen more quickly compared to non-words with low phonotactic probability. However for children with LI, Gray and colleagues (2014) suggested that because of the deficits in their cognitive processing skills, novel words, regardless of their phonotactic probability, are held in the working memory for only a short period of time and less accurately, leading to poor encoding and storage.

Research has shown that words with low neighbourhood density are learnt more quickly and easily than words with high neighbourhood density (Garlock, Walley & Mesala, 2001). Similarly, Gray and colleagues (2014) found that in their study, for both word identification and word production, typically developing children benefited from a low neighbourhood density effect. The authors suggested that because novel words with low-neighbourhood density are similar to only a few known words, the trigger for word learning occurs much faster in typically developing children (Storkel & Adlof, 2009). Additionally, as there is less competition of similar words in the language, the cognitive-processing demands for phonological and semantic activation to learn the new word are reduced (i.e. less inhibition takes place), facilitating the learning of the novel word (Garlock et al., 2001; Gray et al., 2014). However, for preschool children with LI, a high neighbourhood density advantage is observed instead (Gray et al., 2014). Gray and colleagues (2014) proposed that due to deficits in processing capacity and skills, children with LI create and store novel words as a whole rather than creating and storing sub-lexical representations (phonological sequence) of the words. A novel word with high neighbourhood density is similar to many known words. When a novel word with high neighbour density is presented, activation of known words with similar

neighbourhood density occurs. The activation of familiar words assists children with LI to retain the novel word in their working memory while learning takes place (i.e. encoding and storage of the novel word).

The above review of the literature suggests that to maximise the difference in word learning performance between children with and without LI in a DA context, non-words have to be of high phonotactic probability and low neighbourhood density.

5.4.2 Number of Presentations

The number of presentations can influence word learning performance in preschool children with and without LI (Kan & Windsor, 2010, Kapantzoglou et al., 2013). To maximise the word learning performance of preschool children with and without LI, the number of presentations needs to be carefully considered to avoid both floor and ceiling effects (Kapantzoglou et al., 2012). This is so that there is still sufficient support for typically developing children to reveal their word learning ability/potential while challenging to children with LI to reveal their processing limitations.

The number of presentations of each novel word in past word learning studies has been variable. Some studies have investigated children's word learning skills using one or two presentations for each target word (i.e. Dollaghan, 1987) whereas, in other studies (i.e. Gray, 2003), each target word was presented more than 10 times. Rice and colleagues (1994) investigated the performance of 60 5-year-old preschool children with and without LI (30 in each group) on a novel word learning task. Eight novel words were presented in a quick incidental learning context (QUIL) under the conditions of low (each word presented three times) or high exposure (each word presented 10 times). The words were embedded and presented in a video story presentation. Under the condition of low exposure, group differences in the total number of words comprehended were observed. However, under the condition of high exposure, children

with LI were able to comprehend the same number of words as their typically developing counterparts. This result suggests that with increased exposure, word learning ability in children with LI can be facilitated.

Similarly, Gray (2003) examined the performance of 4- to 6-year-old preschool children with and without LI on a word learning task using a trials to criterion approach (i.e. number of trials to reach predetermined learning criterions for comprehension and production of the set of 8 words) and found that typically developing children produced and comprehended more words in fewer trials in comparison with children with LI. Gray concluded that children with LI needed twice as much exposure to the novels words as their peers before being able to comprehend and produce them independently.

The results from both studies suggest that children with LI have limited cognitive processing ability to learn novel words under the condition of low/limited exposure, but with increased exposure to novel words their limitations in cognitive processing may be compensated for to facilitate word learning. Thus the number of presentations in word learning task should remain low so that the word learning ability in typically developing children can be facilitated to reveal their word learning potential while the task remains challenging enough for children with LI to reveal their processing limitations.

Similar to the results from monolingual studies, the difference in word learning ability between bilingual preschool children with and without LI is also maximised under the conditions of low exposure. Kapantzoglou and colleagues (2012) examined the word learning ability of 4- to 5-year-old predominantly Spanish-dominant bilingual children with and without LI on a word learning task (three novel words) in a DA context but in three different exposure conditions (nine, 18 and 27 exposures). Group difference in word identification was only recorded under the condition of nine exposures, and classification accuracy (Sensitivity: 76.9%; Specificity:

80.0%) combined with examiners' score of children's modifiability was highest under the condition of nine exposures. However, under the condition of 18 and 27 exposure, group difference in word identification was not significant.

The results from the studies suggest that to maximise the difference in word learning performance between children with and without LI, the number of exposure to each novel word in a DA context should be low (i.e. nine or fewer exposures). This is so that the word learning task remains challenging for children with LI to reveal their processing limitations yet provide sufficient support to typically developing children to reveal their language learning potential.

5.5 Chapter Summary

The current study has adopted the Validity Argument Framework (Kane 2006; 2013) to gather and evaluate evidence at each stage of the DA development to support the valid interpretation and use of the proposed DA process for identifying bilingual children at risk of LI. The first step in developing a DA that can assist clinician/SLPs in validly identifying bilingual children at risk of LI is to select a developmentally appropriate language task to be evaluated in the DA context.

In this chapter, the evidence to support the first claim (*Observations of performance on the DA can reveal bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI*) under the *Domain Definition* of the Validity Argument Framework (Kane 2006; 2013) was gathered and evaluated. A review of the literature provided unfavourable evidence for using poor performance in grammatical-type tasks (i.e. poor verb tense markings) under linguistically-based accounts to identify bilingual children at risk of LI. The acquisition and mastery of verb tense marking among bilingual children are variable and influenced by the amount of input/exposure they have had to the language. Typically developing young bilingual children who are in the early stages of learning at least two

languages may also display grammatical error patterns that are observed in monolingual children with LI.

On the other hand, there is favourable evidence for using poor performance in language processing tasks under the processing-based account to identify bilingual children at risk of LI. Language processing tasks evaluate children's underlying cognitive ability to process new language information. This is parallel to the purpose of DA, which is to evaluate children's language learning potential. To consider the processing limitations observed in children with LI and to extend it to a language task which can appropriately evaluate early language ability in young bilingual children (i.e. vocabulary acquisition), a word learning task was identified. The review of the literature provides emerging evidence that both monolingual and bilingual children with LI exhibit poor word learning ability compared to their age-matched typically developing peers. This demonstrates that poor performance in a word learning task can be a potential behavioural indicator for LI. It is, therefore, appropriate to evaluate children's word learning ability in a DA context to identify bilingual children at risk of LI. Additionally, it is important to consider the task characteristics of word learning tasks such as word type and number of presentations as they have been found to impact on the word learning performances of children with and without LI.

To recap, the first step in the argument, *Domain definition*, is supported to make the first claim: Observations of performance on the DA (word learning task) can reveal bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI (refer to Table 5-1). The favourable evidence for this claim comes from the evidence identified in the review of the literature where behavioural indicators of LI under each theoretical account of LI were evaluated and considered. The evidence favoured the selection of a word learning task to be evaluated in the DA, to identify bilingual children at risk of LI. In the next chapter, the development of a DA of word learning skills will be presented.

Table 5-1. Evidence to support the claim under the Inference Domain Definition.

Inference	**Claims**	**Assumptions**	**Evidence**
Domain Definition	Observations of performance on the DA can reveal bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI.	The language task selected to be evaluated in the DA is developmentally appropriate and least biased for evaluating language learning potential among bilingual preschool children.	A review of the literature identified language processing tasks under the processing-based account of LI to evaluate the language learning potential of bilingual children with least biased. This is because language processing tasks evaluate children's ability to learn new language materials where their performance is least influenced by their prior language experience. Word learning is identified as a developmentally appropriate language processing task to evaluate the language learning potential of young bilingual children as vocabulary acquisition is one of the first sublinguistic domains that young children acquire. A review of the literature provided emerging evidence that preschool children with LI performed poorly on word learning tasks compared to their typically developing counterparts under low exposure conditions. Therefore poor performance on word learning tasks under low exposure conditions can be a behavioural indicator for LI and identify bilingual children at risk of LI.

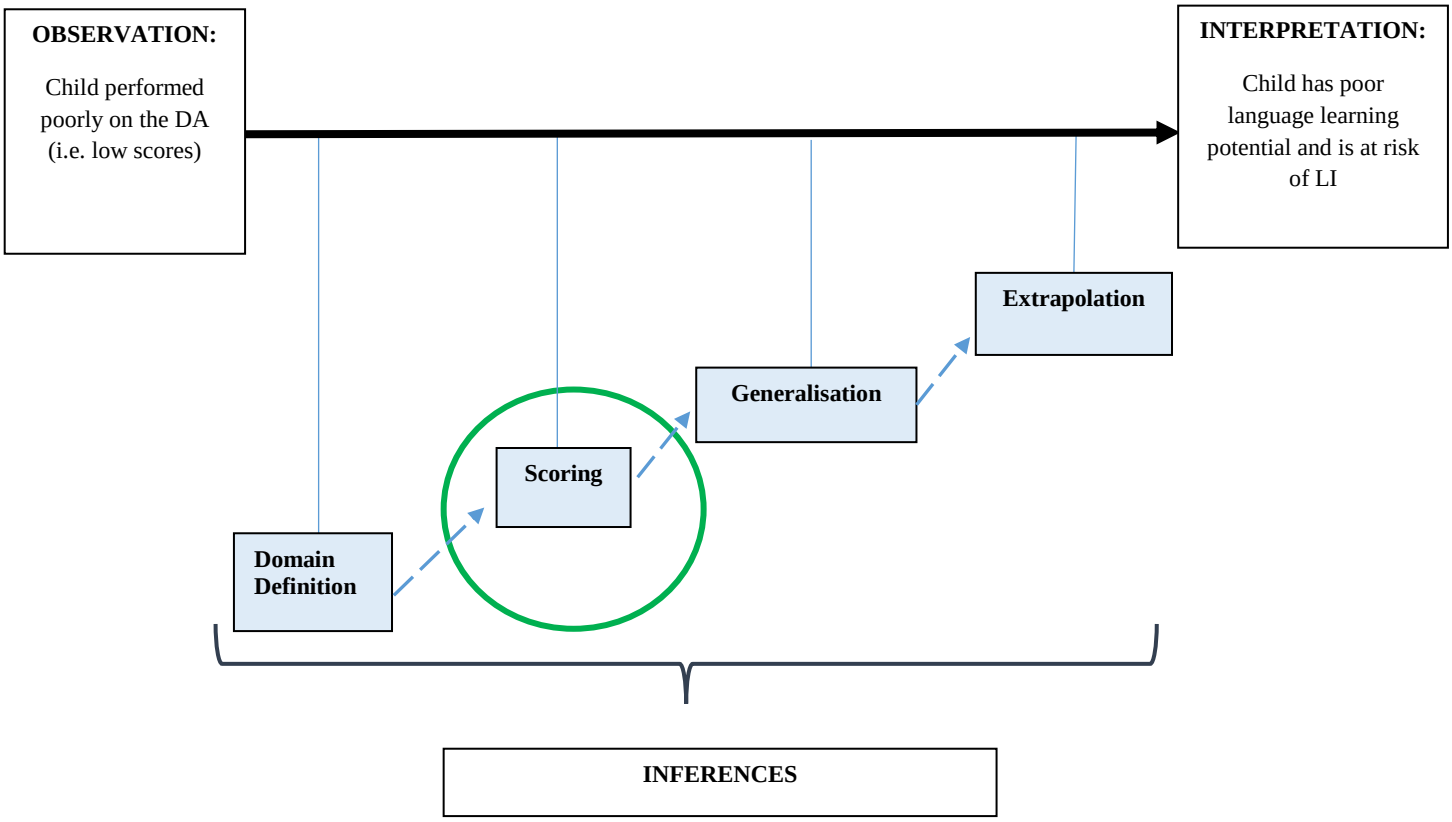
Chapter 6 Phase Two: Development of the Dynamic Assessment of Word Learning Skills

6.1 Introduction

In the preceding chapter (Chapter 5), the evidence supporting the selection of a word learning task to evaluate the language learning potential of bilingual preschool children in a DA context was presented. In this chapter, the development of the DA of word learning skills is presented. The main components of the DA of word learning skills will be highlighted, and the assessment procedure (i.e. administration, scoring) will be described with reference to the supporting literature presented in Chapter 4 and Chapter 5.

The chapter is divided into two main sections. In the first section (6.2), the key design components of the DA of word learning skills will be highlighted. This section focuses on the development and purpose of each component in relation to the evidence from the literature presented in Chapter 4 and 5. In the second section (6.3), a pilot study that was conducted to trial the administrative and scoring procedure of the DA of word learning skills is presented. The purpose of the pilot study was to also gather evidence to support the second claim (*Observations of performance on the DA are evaluated to provided observed scores reflective of language learning potential)* under the Validity Argument Framework (Kane 2006, 2013) regarding the relevance and accuracy of the manner used to evaluate the performance of bilingual children's word learning performance on the DA (circled in green in Figure 6-1). This is also to ensure that the bilingual children's word learning performance on the DA of word learning skills are accurate summaries of the word learning ability consistent with their language developmental profiles. The revisions that resulted from the pilot study will also be presented in this section.

Figure 6-1. Validity Argument Framework – Scoring.



6.2 Key Design Components of the Dynamic Assessment of Word Learning Skills

6.2.1 Length of Assessment

In the Phase 1 survey, length of assessment was listed as one of the top three factors to consider by SLPs when designing a localised assessment approach/tool (see Chapter 3, Section 3.3.5). Therefore, to increase the clinical acceptability of DA as an alternative assessment approach for assessing the language skills of bilingual children, it was crucial that the total time to administer and score the DA was achievable within a single clinical session (i.e. 45 minutes or less).

The DA of word learning skills was designed to be completed in a single assessment session of 30-45 minutes. Preschool children have shorter attention spans compared to school-aged children hence a short and single assessment session was deemed more developmentally

appropriate. The length of administrative time for the DA of word learning skills was also comparable to the length of assessment time required of commercially-available assessment tools to evaluate the language skills of preschool children (i.e. *CELF P2*: Semel, Wiig & Secord, 2004; *PLS 4*: Zimmerman, Pond & Steiner, 2002).

6.2.2 Target Language

The DA of word learning skills was designed to be conducted in the English language as English is commonly an official language spoken in predominantly bilingual countries (i.e. Singapore, Hong Kong, India, Malaysia, and Canada). Children in many predominantly bilingual countries are required to learn the English language and another language (i.e. usually their mother tongue).

6.2.3 Stimuli

A list of non-words as opposed to unfamiliar English words was selected as target words to be learnt in the DA of word learning skills. This was to minimise any potential test bias that could be associated with prior language exposure and to avoid the possibility that any word learning potential observed on the DA of word learning skills was a result of the children's true word learning ability. Another advantage of choosing non-words was that word characteristics such as phonotactic probability and neighbourhood density could be controlled. Phonotactic probability and neighbourhood density not only affect word learning performance but are also known to have different effects on the word learning performance in children with different language developmental profiles (Gray et al., 2014; Kapantzolgou et al., 2012; Storkel & Lee, 2011).

Non-words with high phonotactic probability and low neighbourhood density were chosen as target words to maximize the difference in word learning performance between typically developing bilingual preschool children and bilingual preschool children with LI on the DA.

The evidence and justification for controlling word characteristics of novel words in word learning tasks and its effects on the word learning performance of children with and without LI were previously presented in Chapter 5 (Section 5.4.1). To briefly recap, non-words with high phonotactic probability are known to facilitate word learning in typically developing preschool children whereas the properties of phonotactic probability have no effects on word learning performance in preschool children with LI (Gray et al., 2014; Storkel, 2001; 2003). On the other hand, non-words with low neighbourhood density are known to facilitate word learning in typically developing preschool children while non-words with high neighbourhood density are known to facilitate word learning in preschool children with LI (Garlock et al., 2001; Gray et al., 2014; Storkel & Adlof, 2009).

A literature search was conducted by the researcher to identify suitable non-words that met the criteria for word characteristics (i.e. non-words must be of high phonotactic probability and low neighbourhood density) to be controlled. The researcher eventually selected four monosyllabic non-words of high phonotactic probability and low neighbourhood density used in Storkel et al. (2006) and Gray et al. (2014) studies as the target words to be learnt in the DA of word learning skills.

These four target words were CVC (consonant-vowel-consonant) non-words that followed English phonological rules (Table 6-1). CVC structured words were considered most appropriate as it is a frequently occurring phonotactic sequence in the English language and is one of the first few phonotactic sequences young children are exposed to. Moreover, simple non-words compared to complex non-words are preferred when assessing the word learning ability of bilingual children as they are less affected by the variability in language experience among bilingual children (Thordardottir & Brandeker, 2012).

The four target words were each paired with an unfamiliar target object. Each unfamiliar target object acted as a referent object for each target word. Each of the four target objects (unidentifiable seeds and grains, unidentifiable headgear, an unidentifiable set of rings and an unidentifiable rubber object) were also distinguished into four semantic categories (food, clothing, toys and furniture) that should be familiar to preschool children. It was important to differentiate the sematic categories as the semantic functions would be used to describe the category and/or function of each target object during the 'Teach' Phase. This would assist the children in learning the target words meaningfully and in context.

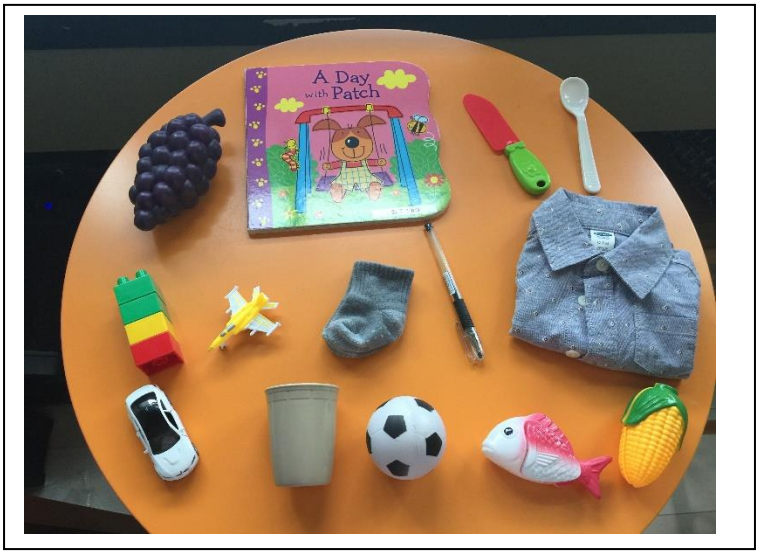
Table 6-1. The four target words (CVC non-words) and their paired target objects.

Target words	Target objects	Associated semantic category
/nɛp/		Furniture
/hɑn/		Food
/jɪb/		Toy
/paib/		Clothing

Another 14 monosyllabic familiar words in English (nouns: grape, book, knife, spoon, blocks, plane, sock, pen, shirt, car, cup, ball, fish, corn) and their referent objects were chosen as

unrelated/distraction stimuli (Figure 6.2). These 14 familiar words and their objects were presented with the target words during the 'Teach' and 'Retest' phase. The role of these 14 familiar words in the DA of word learning skills will be further elaborated on below (Section 6.2.3).

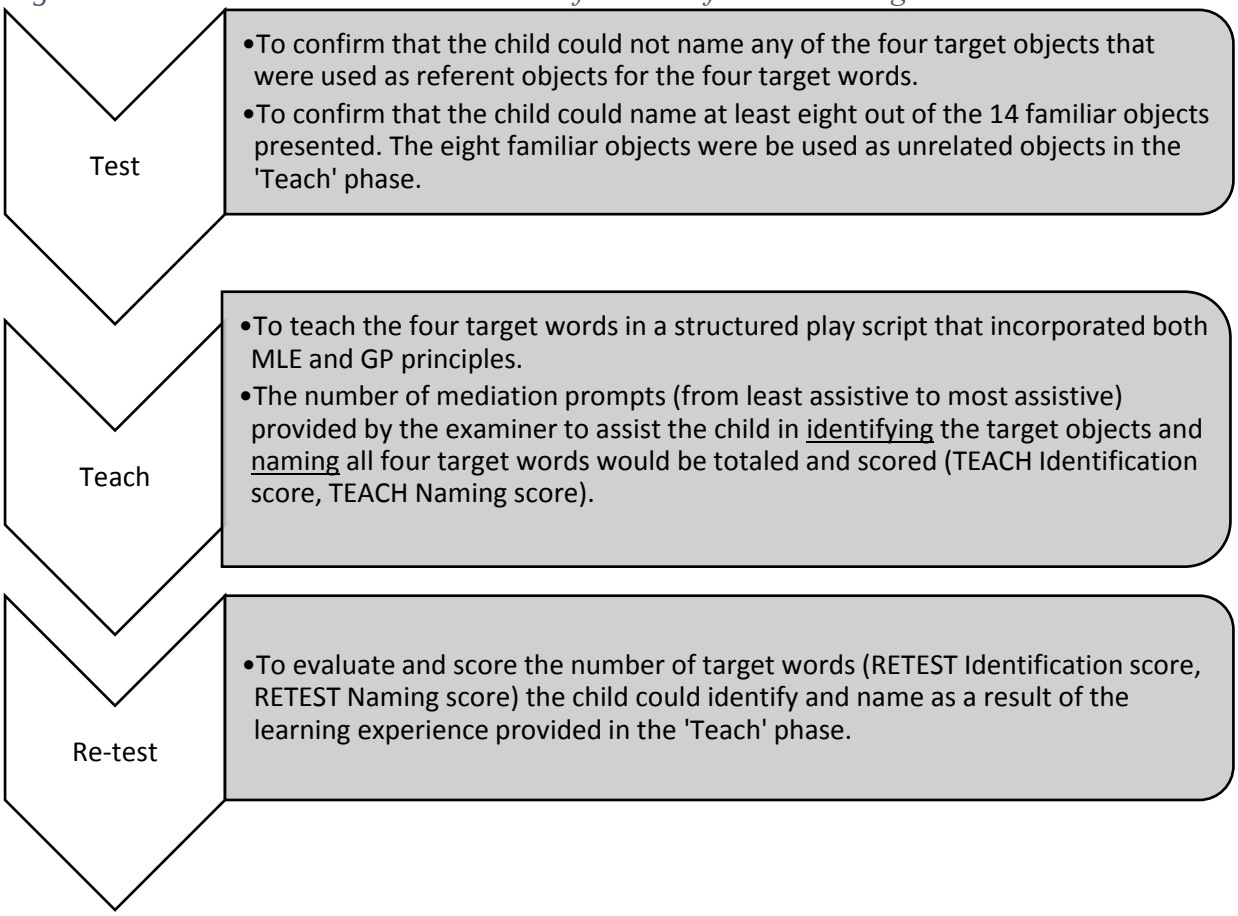
Figure 6-2. Familiar objects used in the DA of word learning skills.



6.2.4 Test-Teach-Retest Format

A 'Test-Teach-Retest' model was selected for the DA of word learning skills as it was found to be an appropriate paradigm to evaluate the language learning potential of bilingual children (Gutierrez-Clellen & Pena, 2001). Specifically, as elaborated in Chapter 4 (Section 4.3), the 'Test-Teach-Retest' model has previously been used in a DA of word learning skills (Kapantzoglou et al., 2012) and a DA of vocabulary (Hasson et al., 2013) to differentiate between the performance of bilingual children referred for speech and language therapy/at risk of LI and typically developing bilingual children. Figure 6-3 below illustrates a general overview of the 'Test-Teach-Retest' format in the DA of word learning skills that was developed in this study.

Figure 6-3. The 'Test-Teach-Retest' model of the DA of word learning skills.



The purpose of the 'Test' phase was to confirm that the child (examinee) was unable to name any of the four target objects. This testing was necessary to ensure that the four target objects were appropriate to be used as referent objects for the four target words before any teaching commenced (i.e. TEST score should be 0). The examiner also had to confirm that the child could name at least eight out of the 14 familiar objects presented. Eight familiar objects would be used as unrelated objects (two at a time) to be presented with each target object during the 'Teach' phase (i.e. two familiar objects and one unfamiliar target object).

In the 'Teach' phase, the target words were taught in an interactive play activity with the incorporation of MLE and GP principles as teaching instructions (to be elaborated in Section 6.2.5). The teaching instructions provided in the play activity were pre-determined and scripted. Mediation prompts, from least to most assistive, were provided to assist the child in identifying all four target objects and naming all four target words (to be elaborated in Section 6.2.5). The

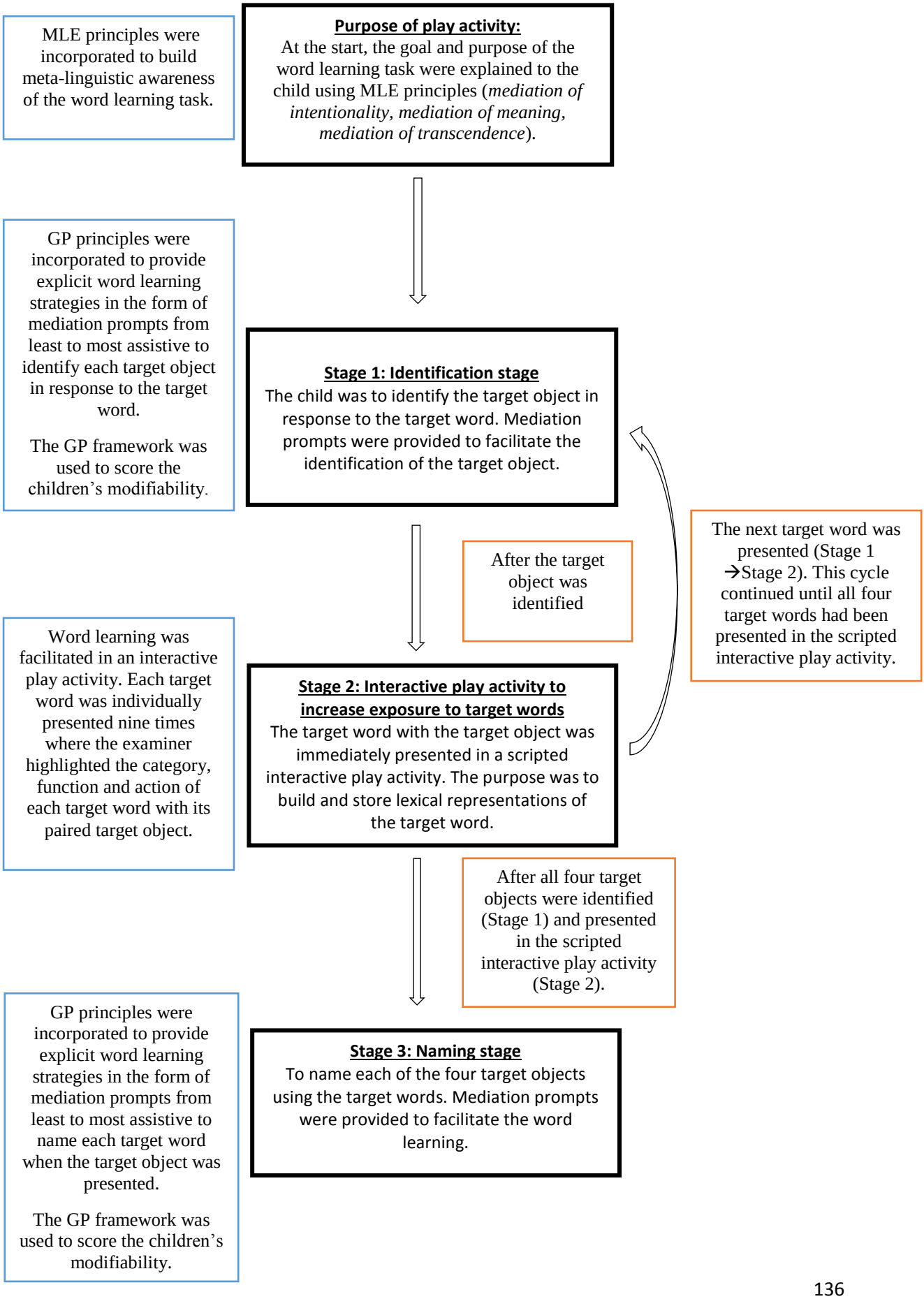
child's modifiability was evaluated by observing and scoring the total number of mediation prompts provided by the examiner to: 1) identify the target objects in response to the target words (i.e. TEACH Identification score) and, 2) name all four target words when each of the target object was presented (i.e. TEACH Naming score). It was hypothesised that bilingual children with LI, due to limitations in their cognitive-processing ability based on the processing-based account of LI, would require more mediation prompts in comparison to their typically developing peers in identifying all four target objects and naming all four target words.

In the 'Retest' phase, the examiner evaluated and scored the number of target words the child could name and identify independently (RETEST Identification score, RETEST naming score) as a result of the learning experience provided for all four of the target words in the 'Teach' phase. It was hypothesised that bilingual children with LI, due to their limitations in their cognitive-processing ability based on the processing-based account of LI, would identify fewer target objects and name fewer target words compared to their typically developing peers.

6.2.5 Teaching Instructions

The structured teaching instructions provided in the 'Teach' phase incorporated both MLE (Lidz, 1991) and GP (Campione & Brown, 1987) principles, while a GP framework was adopted to score the children's modifiability. The learning instructions provided at each stage of the 'Teach phase' are illustrated in Figure 6-4 (blue text box). In-depth discussion of both MLE and GP principles were previously presented in Chapter 4 (Section 4.3).

Figure 6-4. Teaching instructions embedded in the 'Teach' phase.



To briefly recap, MLE teaching instructions have the potential to create meaningful environmental and socialization experiences to support language learning. However, MLE teaching instructions can be implicit and challenging for young preschool children, whose metacognition/metalinguistic skills are still developing, to comprehend and follow (Haywood and Lidz, 2006). This is especially true for bilingual preschool children who are non-English dominant. Moreover, the provision of MLE teaching instructions is often unscripted and relies heavily on language as mode of delivery (Camilleri & Bottling, 2013). In order to address these limitations and to create a meaningful and purposeful learning experience, the MLE teaching instructions were simplified, scripted and embedded in a play script in the DA of word learning skills.

However, bilingual preschool children whose metacognitive/metalinguistic skills are still emerging may not fully benefit from MLE teaching instructions that require them to apply cognitive-linguistic skills (i.e. analytical perception, analogical reasoning) to develop their own word learning strategies. Young bilingual preschool children would benefit more from specific word learning strategies that are explicit and directive instead. The incorporation of GP principles into the teaching instructions allowed the provision of explicit word learning strategies (i.e. semantic or phonological cues) to facilitate word learning. Additionally, the GP framework also allowed examiners to score children's modifiability objectively (i.e. based on the level of mediation prompts provided).

In the subsections below, the teaching instructions provided at each stage of the 'Teach' phase are presented.

6.2.5.1 Teach Phase: Purpose of the Play Activity

The word learning task was designed as a scripted interactive play activity with a teddy bear (henceforth referred to as Teddy) where Teddy had some objects that he would like to show to the child. Four unfamiliar target objects and eight familiar objects (the eight familiar objects that the child had earlier named in the 'Test' phase) were used as stimuli. The child was provided with opportunities in the play activity with Teddy to learn the novel names of the four target objects.

The examiner first established the child's metalinguistic awareness of the word learning task by stating the goal and purpose of the activity. The examiner then read out a pre-determined script where MLE principles (i.e. *mediation of intentionality, mediation of meaning, mediation of transcendence, mediation of competence*) had been incorporated to establish the child's awareness to the goal and purpose of the activity. At the start, the goal of the activity was stated explicitly (*mediation of intentionality*) by the examiner, for example, "*Today we are going to play and learn the special names of the things we see… Teddy here is going to help."* Next, the purpose of the activity was communicated, for example, *"Why do we need to learn special names? Special names help us tell things apart…"* This was followed by connecting the activity with the child's previous experiences *(mediation of transcendence)*, for example, "*Do you remember the special names of these things I showed you just now? Wow, you know their special names*..." Finally, the child was told of the plan for the activity *(mediation of competence)*, for example, "*Teddy is going to show you some things… Teddy will teach you the names of the things you don't know and you will try your best to remember the names*." The full script is presented in Chapter 7.

6.2.5.2 Stage 1: Identification

Once the examiner had established the goal and purpose of the activity for the child using MLE principles, the examiner proceeded to ask the child to identify the target object in response to

target word presented. For each target object to be identified, the child was presented with three objects on a table with Teddy. Two were familiar objects that the child had previously named in 'Test' phase, and one an unfamiliar object (referent object for the target word). The examiner then presented the target word by asking, *"Teddy asks which one is /XXX/."* The child had to point to the object he/she thought the target word was referring to. Mediation prompts from the least to most assistive were provided systematically to support the child in identifying the target object correctly. Table 6-2 below describes the three levels of mediation prompts that were provided for the identification of the target object in response to the target word.

The mediation prompts provided were modified from a hierarchy of cues used in a vocabulary learning task by Camilleri and Law (2007) to support children in the identification of target pictures in response to the target vocabulary words. The hierarchy of mediation prompts is designed to allow the child to be an active learner in establishing new word-unfamiliar object matches (Camilleri & Law, 2007). The child is to use the principles of 'relevance, 'discrepancy' (Bloom, 2000, cited in Camilleri & Law, 2007, p. 315) and 'mutual exclusivity' (Clark, 1993, cited in Camilleri & Law, 2007, p. 315) in identifying the target object among familiar objects.

Table 6-2. The levels of mediation prompts provided to identify each target object in response to the target word.

Mediation prompts (least assistive to most assistive)	**Description**	**Example**
Level 1: Contextual mediation	The examiner first placed two familiar objects and one unfamiliar object (referent object for target word) on the table. The target word was then presented by the examiner. The child was to point to the object that he/she thought referred to the target word. a) If the child was able to identify the target	Examiner: "*Teddy asks which one is /han/?"*

	object correctly, no further mediation prompts was provided. The examiner would proceed to Stage 2 of the 'Teach' Phase.	The child correctly points to the unfamiliar target object. Stage 2 of the 'Teach' Phase proceeds, for example, Examiner: "*Yes! This is /han/…"*
	b) If the child was unable to identify the target object, the examiner provided feedback for the incorrect match before proceeding to provide the next level of mediation prompt.	The child points to a familiar object instead. Examiner: *"No, that is not /han/. That is a hard word. Let's find the one you know first."*
Level 2: Implicit contextual and language mediation	Next, the examiner asked the child to identify the two familiar objects on the table.	Examiner: "*Teddy asks which one is ball (familiar object)?"* – waits for the child to point. Examiner: *"Teddy asks which is spoon (familiar object)?"* – waits for the child to point.
	After the child pointed to both familiar objects, the target word was presented again.	"*Now Teddy asks which one is /han/?*"
	a) If the child was able to identify the target object correctly, no further mediation prompts was provided. The examiner proceeded to Stage 2 of the 'Teach' Phase.	The child correctly points to the unfamiliar target object. Stage 2 of the 'Teach' Phase proceeds, for example, Examiner: "*Yes! This is /han/…"*
	b) If the child was unable to identify the target object, the examiner provided feedback for the incorrect match before proceeding to provide the final level of mediation prompt.	The child points to a familiar object instead. Examiner: *"No, that is not /han/. Let us see what we have here. "*

Level 3: Explicit contextual and language mediation	Next, the examiner proceeded to name all three objects while pointing to each. The target word was presented again. Explicit identification of the target object in response to the target name had occurred. Therefore, the child was expected to be able to identify the target object correctly.	Examiner places Teddy behind each object while naming each object: "*This is 'ball' (familiar object), this is 'spoon' (familiar object), this is /han/ (unfamiliar target object).*" *Examiner: "Now Teddy asks Which one is /han/?"* Child points to target object successfully. Stage 2 of the 'Teach' Phase proceeds, for example, Examiner: "*Yes! This is /han/…*"

For each target object, the examiner recorded the level of mediation prompt required by the child to identify the target object in response to the target word. If only the first level of mediation prompt was provided, one point was assigned. If up to two levels of mediation prompts (Level 1 and Level 2) were provided, two points were assigned. If up to three levels of mediation prompts (Level 1, Level 2 and Level 3) were provided, three points were assigned. The total points required to identify all four target objects added up to the child's overall 'TEACH Identification score'.

It was hypothesised that bilingual children with LI, due to limitations in processing capacities/skills based on the processing-based account of LI, would require more mediation prompts compared to their typically developing peers in identifying all four target objects correctly.

6.2.5.3 Stage 2: Interactive Play Activity to Increase Exposure to Target Words

Once the child had identify the target object (Stage 1), the target word and the target object were immediately presented in a scripted interactive play activity with the Teddy. The purpose was to increase exposure and facilitate word learning for each target word in an interactive play setting. The examiner provided support for learning each target word by 1) stating the semantic category to which the target object belonged (i.e. toy, food), 2) describing the function of the target object while providing a corresponding gestural action (i.e. Teddy can eat it – shows teddy bear eating the object), 3) allowing the child to hold and manipulate the target object with Teddy, and 4) providing opportunities for the child to imitate/use the target word during the play (Kapantzoglou et al., 2012). At the end of the play script for each target word, the examiner commented on the effort made by the child by stating: *"You have learnt a special name for this thing."* The examiner also reminded the child the purpose of the activity by saying: *"Remember, you have to try your best to remember the special name."*

To maximise the difference in word learning performance between bilingual preschool children with and without LI on the DA of word learning skills, the number of presentations of each target word in the play script was carefully considered. This was so that there was sufficient teaching support for typically developing children to reveal their word learning ability/potential while challenging children with LI to reveal their processing limitations (Kapantzoglou et al., 2012). A review of the literature was presented in Chapter 5 (under Section 5.4.2). To briefly recap, the difference in word learning performance in preschool children with and without LI was maximised under the condition of low exposure (i.e. limited number of presentations) (Gray, 2003; Kapantzoglou et al., 2012; Rice et al., 1994). For instance, Kapantzoglou and colleagues (2012) investigated the word learning performance of preschool bilingual children with and without LI in DA context under the conditions of different levels of exposure (Kapantzoglou et al., 2012). Group difference in word learning performance was only

significant under the condition of low exposure (i.e. nine presentations) but not under the condition of higher exposures (i.e. 18 presentations, 27 presentations).

In the DA of word learning skills, each word was presented nine times in the interactive play script to facilitate word learning in typically developing bilingual preschool children. It was hypothesised that under the condition of low exposure, the word learning task would still be supportive for typically developing children to reveal their word learning ability/potential while challenging to children with LI to reveal their processing limitations.

6.2.5.4 Stage 3: Naming

After all four target words and their target objects were presented individually in the interactive play activity (Stage 2) (refer to Figure 6-4), the third and final stage in the 'Teach' phase was presented. At this stage, each target object was once again presented on the table with Teddy. The examiner then encouraged the child to name the target object by saying, "*Teddy asks what is this?*" Mediation prompts from the least to most assistive were provided systematically to support the child in naming the target object correctly. Table 6-3 below describes the four levels of mediation prompts that were provided to support the child in naming each target object with the target word.

The mediation prompts provided were modified from the hierarchy of cues used in a word learning study by Burton and Watkin's (2007) to support children in eliciting the target words when their respective referent objects were presented. The hierarchy of prompts included both semantic and phonological prompts to facilitate word retrieval in children (McGregor, 1994; Wing, 1990).

Table 6-3. The levels of mediation prompts provided for naming each target object with the target word.

Mediation prompts (least assistive to most assistive)	**Description**	**Example**
Level 1: Elicitation question	The examiner placed one target object on the table, followed by asking the child to name the object.	The examiner places an unfamiliar target object on the table (i.e. target object for */han/.* Examiner: *"Teddy asks what is this?"*
	The child attempted the name the target object. a) If the child was able to name the target object correctly, no further mediation prompts was provided.	Child names the target object correctly. The examiner praises the child. Examiner: *"Wow you have just learned a new name..."*
	b) If the child was unable to name the target object correctly, the examiner provided the feedback, and the next level of mediation prompt was provided.	The child is unable to recall the target word or names the object incorrectly. Examiner: *"No, this is not /XXX/(child's response). Let's try again."*
Level 2: Mediation with semantic prompt	Next, the examiner described the category or function of the target object.	Examiner: *"Teddy likes to eat this. What is this?"*
	The child attempted to name the target object a) If the child was able to name the target object correctly, no further mediation prompts was provided.	Child names the target object correctly. The examiner praises the child. Examiner: *"Wow you have just learned a new name…."*
	b) If the child was unable to name the target object correctly, the	The child is unable to recall the target word or names the target object incorrectly.

	examiner provided the feedback, and the next level of mediation prompt was provided.	Examiner: *"No, this is not /XXX/(child's response). Let's try again."*
Level 3: Mediation with phonological prompt	Next, the examiner provided the starting sound of the target word. The child attempted to the name the target object a) If the child was able to name the target object correctly, no further mediation prompts was provided. b) If the child was unable to name the target object correctly, the examiner provided the feedback, and the final level of mediation prompt was provided.	Examiner: *"The name starts with /h/, What is it?"* Child names the target object correctly. The examiner praises the child. Examiner: *"Wow you have just learned a new name…."* The child is unable to recall the target word or names the object incorrectly. Examiner: *"No, this is not /XXX/ (child's response). Let's try again."*
Level 4: Mediation with model	Lastly, the examiner named the target object. The target word had been named explicitly. Therefore, the child was expected to name the target object correctly.	Examiner: *"This is /han/. What is this?"* Child names target object correctly. The examiner praises the child. Examiner: *"Wow you have just learned a new name…"*

For each target object presented, the examiner recorded the level of mediation prompt required by the child to correctly name it. If the child was able to name the target object correctly with the first level of mediation prompt (Level 1: Elicitation question), it was recorded as one point. For every additional level of mediation prompt provided subsequently, one more point was added. For example, if the final mediation prompt (i.e. Level 4) was provided, a total of four

points was allocated. The number of mediation prompts for each target object to be named correctly was totalled to become the child's overall 'TEACH Naming score'.

It was hypothesised that bilingual children with LI, due to limitations in processing capacities/skills based on the processing-based account of LI, would require more mediation prompts compared to their typically developing peers in naming all four target objects.

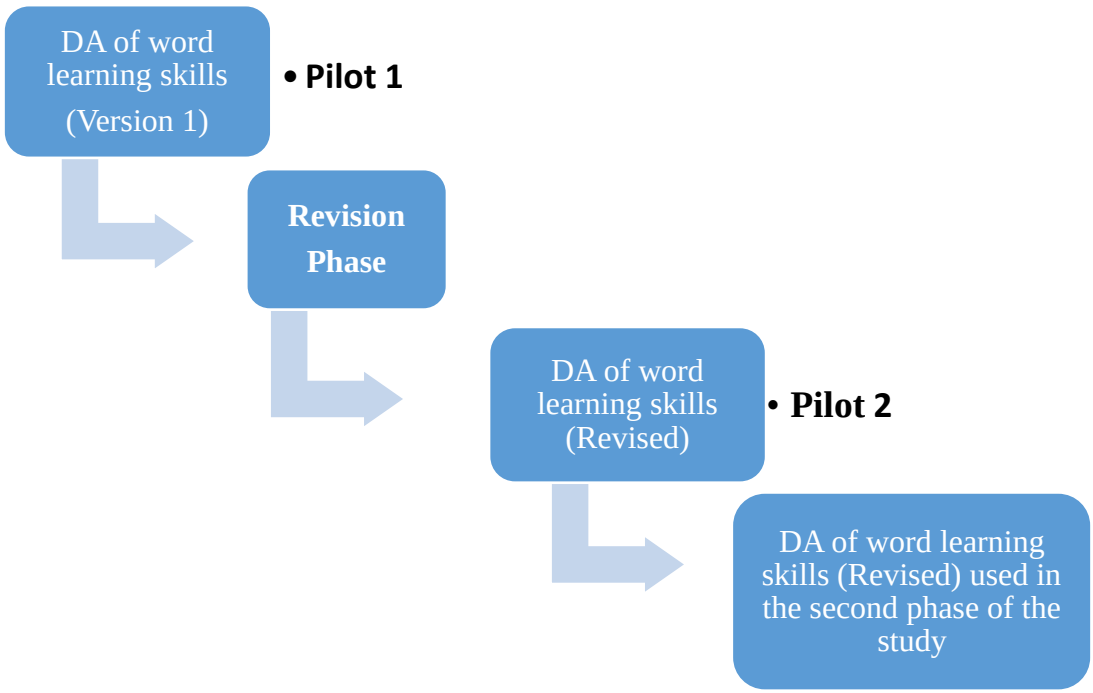
6.3 Pilot Study

Before the DA of word learning skills was used for Phase Two of the study, a small-scale pilot study was conducted. Firstly, this ensured that the key components of the DA of word learning skills (i.e. stimuli, test-teach-retest format, teaching instructions) when put together could be carried out as intended and the DA of word learning skills was a feasible alternative assessment measure (Van Teijlingen & Hundley, 2002). Secondly, the pilot study allowed the review and revision, if required, of the teaching instructions of the DA of word learning skills (Van Teijlingen & Hundley, 2002). This ensured that the teaching instructions provided were developmentally appropriate and could be understood by bilingual preschool children who had at least one year of English exposure in school settings.

Lastly and most importantly, the purpose of the pilot study was also to gather evidence to support the second claim (*Observations of performance on the DA are evaluated to provide observed scores reflective of language learning potential*) under the *Scoring* inference of the Validity Argument Framework (Figure 6-1, pg. 121). This was to ensure that the scoring procedure and scoring rule could be applied objectively and the scores ('Test' and 'Retest' scores) obtained by bilingual preschool children with and without LI were differentiated.

The DA of word learning skills was piloted twice (i.e. Pilot 1, Pilot 2) and revised (i.e. Revision Phase) once before it was used in Phase Two of the study (Figure 6-6).

Figure 6-5. Flow chart illustrating the pilot and revision phases of the DA of word learning skills.



6.3.1 Pilot 1

6.3.1.1 Pilot Participants

Four English-Mandarin bilingual preschool children (aged between 5;9 to 6;5) who were attending local preschools in Singapore participated in Pilot 1. They were sourced from the personal contacts of the researcher. Before participating, verbal consent was obtained from their parents. All testing took place in the children's respective homes, within the line of sight of their parents. Verbal assent from all children were also taken prior to testing.

Two of the children were reported to be typically developing (aged 5;9 and 6;5 respectively) with no reported speech and language concerns from parents. The other two children were reported by their parents to have language difficulties (aged 6;1 and 5;11 respectively) and were receiving speech and language therapy from a private clinic.

6.3.1.2 Procedure

The DA of word learning skills (Version 1) was administered to all four children in their respective home settings.

6.3.1.3 Results

Summary of results:

- All four children were able to name the proposed 14 familiar objects.
- Out of the four target objects, three of them (unidentifiable seeds/grains, unidentifiable rubber object, unidentifiable clothing item) were observed and confirmed to be unfamiliar to all four children. None of the children attempted to name the objects.
- One target object (unidentifiable plastic rings) was named by two children as 'chains' and 'rings' respectively. The chain of plastic rings was deemed partially familiar to some children.
- Two children were observed to have difficulties following the MLE teaching instructions presented at the start of the 'Teach' phase. Some of the MLE teaching instructions were observed to be potentially complex for preschool children to follow and understand.
- The scoring rules could be applied objectively. Scoring was based on 1) the level of mediation prompts provided in the 'Teach' phase for the child to identify and name all four target objects, and 2) the number of target objects identified and named correctly in the 'Retest' phase. Scores could be tabulated on the spot.
- As hypothesised, the children receiving speech and language therapy (i.e. likely children with LI) required more mediation prompts to identify and name all four target objects in the 'Teach' phase compared to the typically developing children (Table 6-4).

- As hypothesised, the children attending receiving speech and language therapy (likely children with LI) identified and named fewer target objects compared to the typically developing children in the 'Retest' phase (Table 6-4).

Table 6-4. Scores obtained by pilot participants in the DA of word learning skills (Version 1).

Participant number	**Pilot 001**	**Pilot 002**	**Pilot 003**	**Pilot 004**
Age	5;9	6;5	6;1	5;11
Language Developmental Profile	Typically developing	Typically developing	Reported language difficulties, receiving speech and language therapy	Reported language difficulties, receiving speech and language therapy
TEACH (Range: 0)	0	0	0	0
TEACH Identification score (Range: 4-12)	4	4	6	5
TEACH Naming score (Range:4- 16)	10	12	14	15
RETEST Identification score (Range: 0-4)	4	4	3	2
RETEST Naming score (Range: 0-4)	2	1	1	0

6.3.1.4 Revisions Made to the DA of Word Learning Skills

The DA of word learning skills was reviewed after Pilot 1, and the following revisions were made:

- An unidentifiable wooden object replaced the set of 'coloured rings' (Table 6-5). The wooden object was also presented to three adults who had extensive experience working with preschool children in Singapore. They agreed that the wooden object was likely an unfamiliar object to most preschool children.
- The MLE teaching instructions were further simplified and shortened (Table 6-5). Three adults who had extensive experience working with bilingual preschool children

were asked to review the revised MLE teaching instructions. All agreed that the MLE teaching instructions could be understood and followed by bilingual preschool children who have had at least one year of exposure to the English language in the school setting.

Table 6-5. Summary of revision made to the DA of word learning skills after Pilot 1.

Components of the DA of word learning skills that were revised	**DA of word learning skills Version 1**	**DA of word learning skills (Revised)**
Stimuli (target object)	Plastic rings that were partially familiar. 	Replaced with an unidentifiable wooden object. 
Teaching instructions	Some teaching instructions were multistep and presented in complex sentences. <u>Example</u> Examiner: "*Today we are going to play and learn the special names of the objects that Teddy is going to show you. I want you to pay attention and try your best to remember them. Teddy here will be helping you if you don't know the names.*"	Teaching instructions were simplified and presented in shorter sentences. <u>Example</u> Examiner: "*Today, we are going to play and learn the special names. You are going to learn the names of some of the things you saw just now. I want you to try your best to remember. Teddy here is going to help you.*"

6.3.2 Pilot 2

6.3.2.1 Pilot Participants

A second group of English-Mandarin bilingual preschool children (*n*=4; aged between 4;11 to 6;3) who were attending local preschools in Singapore participated in Pilot 2. They were also sourced from the personal contacts of the researcher. Similarly, before participating, verbal

consent was obtained from their parents. All testing took place in the children's respective homes, within the line of sight of their parents. Verbal assent from all children were also taken prior to testing.

Two of the children were reported to be typically developing (aged 5;2 and 4;11 respectively) with no reported speech and language concerns from their parents while the other two children were reported to have language difficulties and were receiving speech and language therapy from a private clinic (aged 6;3 and 5;3 respectively).

6.3.2.2 Procedure

The DA of word learning skills (Revised) was administered to all four children (The DA procedure will be presented in Chapter 7).

6.3.2.3 Results

Summary of results:

- All children were able to name all 14 familiar objects.
- All children were unable to identify or name any of the four target objects.
- The revised teaching instructions could be followed and understood by all four children. No repetition of teaching instructions was required.
- The scoring rules could be applied objectively. Scoring was based on 1) the level of mediation prompts provided in the 'Teach' phase for the child to identify and name all four target objects, and 2) the number of target objects identified and named correctly in the 'Retest' phase. Scores could be tabulated on the spot.
- As hypothesised, the children with reported language difficulties required more mediation prompts to identify and name all four target objects in the 'Teach' phase compared to the typically developing children (Table 6-6).

- As hypothesised, the children with reported language difficulties identified and named fewer target objects in the 'Retest' phase compared to the typically developing children (Table 6-6).
- No further revision was required and the DA of word learning skills (Revised) was used in Phase Two of the research study.

Table 6-6. Scores obtained by participants in the DA of word learning skills (Revised).

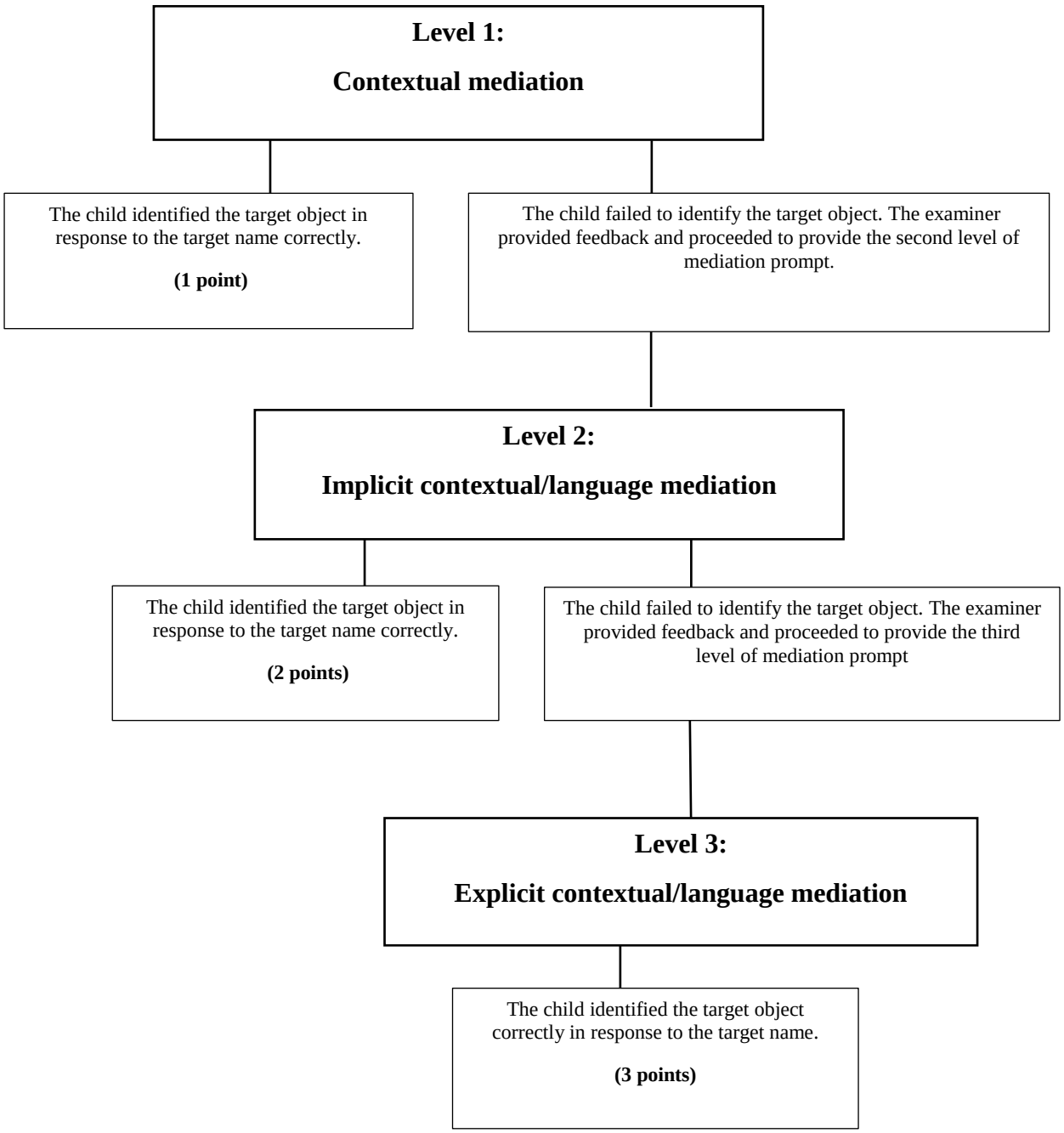
Participant number	**Pilot 005**	**Pilot 006**	**Pilot 007**	**Pilot 008**
Age	5;2	4;11	6;3	5;3
Language Developmental Profile	Typically developing	Typically developing	Reported language difficulties, receiving speech and language therapy	Reported language difficulties, receiving speech and language therapy
TEACH (Range: 0)	0	0	0	0
TEACH Identification score (Range: 4-12)	4	4	4	8
TEACH Naming score (Range:4- 16)	10	12	14	16
RETEST Identification score (Range: 0-4)	4	3	2	2
RETEST Naming score (Range: 0-4)	3	2	0	0

6.3.3 Validity Argument Framework: Evaluating the Evidence under Scoring

In the DA of word learning skills, scoring evaluations occurred at three time points. The first scoring evaluation occurred at Stage 1 of the 'Teach' phase. At this stage, the examiner observed whether the child was able to identify the target object in response to the target word. Depending on the child's initial and subsequent responses, the examiner provided mediation prompts from the least assistive to most assistive (i.e. up to three levels of mediation prompts)

until the child was able to correctly identify each the target object in response to the target word. An illustrated scoring flow chart is presented in Figure 6-7. The scoring for Stage 1 of the 'Teach' phase (i.e. TEACH Identification score) was based on the total number of mediation prompts (Minimum score: 4 to Maximum score: 12) that were provided for the child to identify all four target objects in response to their target names.

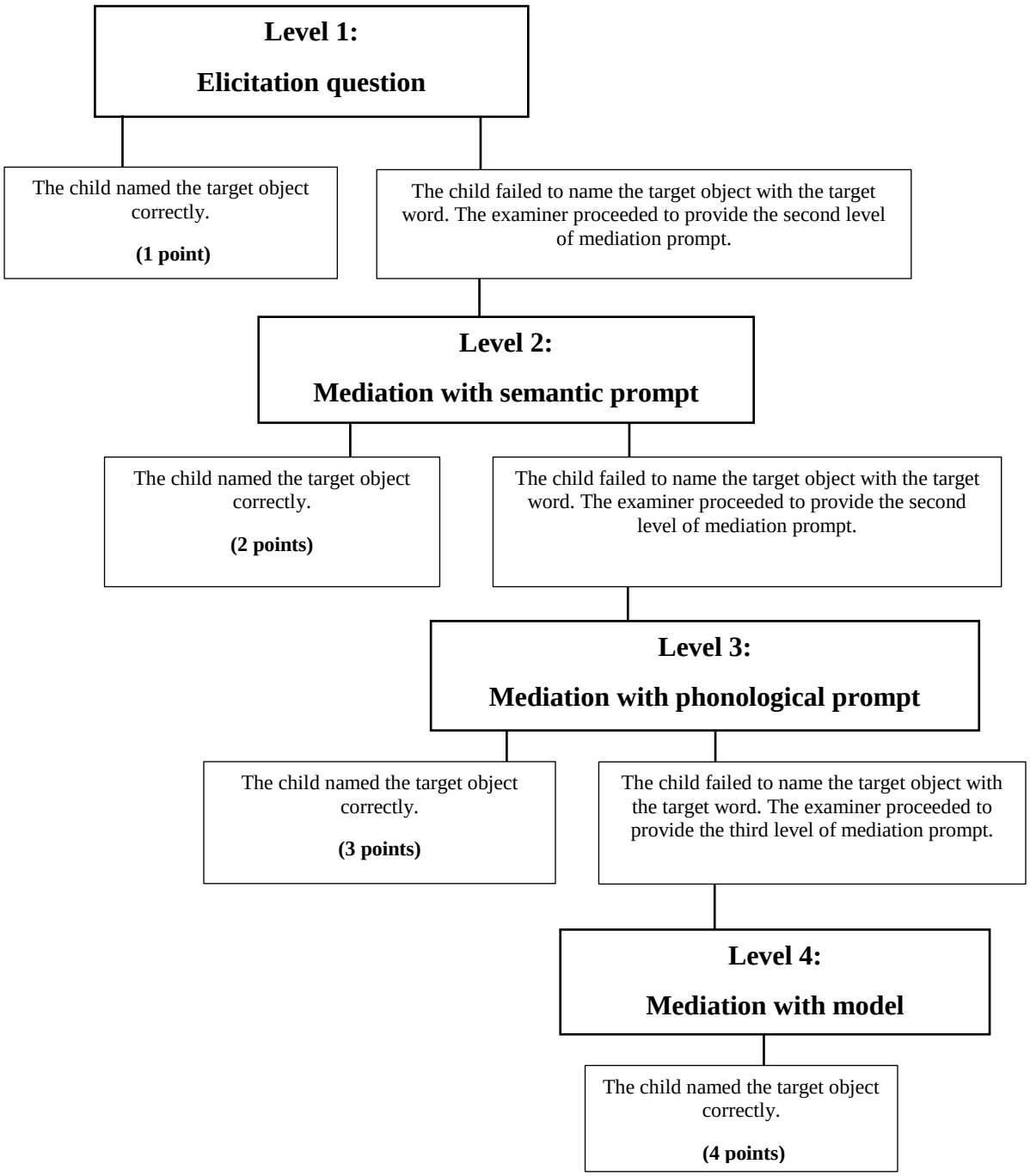
Figure 6-6. Hierarchy of mediation prompts (Identification) and scores.



The second scoring evaluation occurred at Stage 3 of the 'Teach' phase. At this stage, the examiner observed whether the child was able to name all four of the target objects correctly. Each target object was presented individually to the child. Depending on the child's initial and subsequent responses, the examiner provided mediation prompts from the least assistive to most assistive (i.e. up to four levels of mediation prompts for each target object) until the child was able to correctly name each target object with the target word (Table 6-3). An illustrated scoring flow chart is presented in Figure 6-7. The scoring for Stage 3 of the 'Teach' phase (i.e. TEACH Naming score) was based on the total number of mediation prompts (Minimum score: 4 to Maximum score: 16) that were provided for the child to name all four target objects.

The third scoring observation occurred at the 'Retest' phase. At this stage, the examiner observed and recorded the number of target objects that the child could name and identify independently as a result of the teaching opportunities provided in the 'Teach' phase. The examiner set up the naming evaluation by placing all four target objects and four familiar objects (i.e. child had earlier named successfully in the 'Test' phase) on the table. The child was asked to name each object before placing each item into a bag that belonged to Teddy. For each target object that was correctly named, the child was awarded one point. The RETEST Naming score was based on the total number of target objects correctly named. To evaluate the child's ability to identify the target objects when named, all four target objects and four familiar objects were once again placed on the table. The examiner would name an object at random, and the child was to identify the object corresponding to the name. For each target object that was identified correctly in response to the target name, the child was awarded one point. The RETEST Identification score was based on the total number of target objects that were correctly identified in response to their target names when presented.

Figure 6-7. Hierarchy of mediation prompts (TEACH Naming) and scores.



The scoring rubric of the DA of word learning skills was designed with the intention that the examiner's observations of the child's responses corresponded with objective scores. The pilot study confirmed that: 1) the DA procedure provided appropriate opportunities within the assessment itself for word learning to take place as both typically developing children and children with LI were able to learn the target words within their abilities, and 2) the scoring

procedure and scoring rules could be followed and scored accordingly. More specifically, results from the pilot showed that the scoring rules could be applied objectively to accurately differentiate the scores (i.e. based on the level of mediation prompts provided in the 'Teach' phase; number of target words identified or named in the 'Retest' phase) obtained by typically developing bilingual children and bilingual children with reported language difficulties, reflecting their difference in both modifiability and language learning ability/potential. This was expected as the components of the DA of word learning skills were designed to maximise the difference in word learning performance between bilingual children with LI and typically developing bilingual children.

6.4 Chapter Summary

In this chapter, the development of the DA of word learning skills was presented. The key components of the DA of word learning skills including the choice of stimuli (i.e. target words), the 'Test-Teach-Retest' format and teaching instructions (i.e. incorporating both MLE and GP principles) were carefully selected and designed to maximise the difference in word learning performance between bilingual preschool children with LI and typically developing bilingual children in a DA paradigm.

Before the DA of word learning skills was finalised and used in Phase Two of the study, a pilot study was conducted to review the administrative process, scoring procedure and scoring rules. Revisions were made to the DA of word learning skills (i.e. teaching instructions were simplified, an unfamiliar object replaced) to ensure the feasibility of the DA of word learning skills as an alternative assessment approach. Most importantly, the pilot study gathered evidence to support the second claim in the Validity Argument Framework, under the *Scoring* inference (Table 6-7). The scoring procedures and scoring rules were appropriately designed.

The scoring rules could be objectively applied to accurately capture the difference in word learning performance among bilingual children with and without LI. The revised version of the DA of word learning skills that was used in Phase Two of the research study is presented in the next chapter.

Table 6-7. Evidence to support the second claim under the Inference - Scoring.

Inference	**Claims**	**Assumptions**	**Evidence**
Scoring	Observations of performance on the DA task are evaluated to provide observed scores reflective of language learning potential	The task procedure provides opportunities for language learning	Opportunities were created in the 'Teach' Phase for the children to learn all four target words
		The task administrative conditions are appropriate for observing and scoring language learning opportunities	The administrative procedure of the DA was designed to be structured and scripted. This ensured that the learning conditions across children were consistent and optimum for word learning to take place.
		The rubric for scoring responses correspond to the children's language learning performance on the DA task	There was systematic development of the rubric for scoring children's modifiability (i.e. hierarchy of mediated prompts provided in the 'Teach' phase) and language learning potential (i.e. number of target words identified and named correctly in the 'Retest' phase). Results from the pilot study showed that both typically developing children and children with LI were able to demonstrate evidence of language learning. There was a difference in all four scores obtained between typically developing children and children with LI.

Chapter 7 Phase Two: DA of Word Learning Skills

7.1 Introduction

In this chapter, the full procedure for administering the DA of word learning skills is presented (Revised). To illustrate the administrative instructions and components of the DA of word learning skills in detail, the administration of the DA of word learning skills is described in the context of an assessment session between the examiner and a child.

7.2 Test Phase

The DA of word learning skills was administered individually to each child in a quiet setting in a single session. Prior to the administration of the DA of word learning skills, rapport between the examiner and the child was established through a play activity. This ensured that the child was comfortable, enthusiastic and involved for the rest of the session.

The purpose of the 'Test' phase was to confirm that the child could name at least eight familiar objects and none of the unfamiliar target objects. At the start, the examiner began by explaining to the child that he/she would be shown some objects and the child was encouraged to name each object to his/her best ability, for example, "*I am going to show you some things. Some, you may know their names. Some you may not know their names. You will try your best to tell me their names. It is all right if you tell me you don't know their names*." After which, the examiner proceeded to present each object individually (14 familiar objects, four unfamiliar target objects) on the table in random order.

The 'Test' phase ended when the child named at least eight out of the 14 familiar objects and confirmed that he/she was able to name none of the four target objects. The eight familiar objects that were named by the child were presented again in the 'Teach' phase with the four target objects.

7.3 Teach Phase

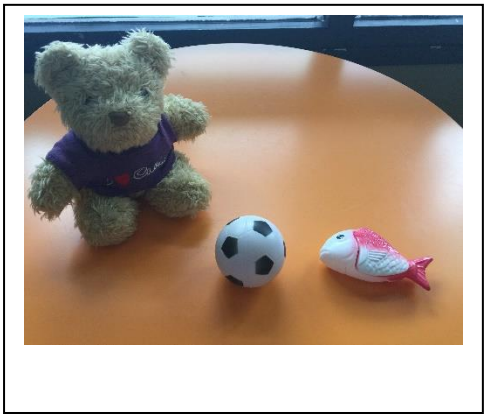
The purpose of the 'Teach' phase was to evaluate the child's modifiability in response to the teaching opportunities provided to learn all four target words. The child's modifiability was measured by the number of mediation prompts provided in Stage 1 and Stage 3 of the 'Teach' phase.

Using the pre-determined script, the examiner began the 'Teach' phase by stating the goal and purpose of the word learning task that was about to occur (**mediation of intentionality**): "*Today we are going to play and learn the special names of the things we see. You are going to learn the names of some of the things you saw just now. I want you to try your best to remember their names. Teddy here is going to help you (examiner puts Teddy on the table).*" (Figure 7-1).

Figure 7-1. Teddy being introduced to the child.

Next, the purpose of the activity (**mediation of meaning**) was explained by relating the activity to the child's previous experiences (**mediation of transcendence**): *"Why do we need to learn special names? Special names help tell things apart. Like how you use special names to tell me the names of the things I showed you just now. We use special names all the time at home and in school."* The examiner then proceeded to place two familiar objects on the table (Figure 7-2) that the child had earlier named in the 'Test phase' and asked: *"Do you know the special names of these things?"* Given that the two familiar objects were earlier named by the child, the child was expected to name both objects easily.

Figure 7-2. Teddy and two familiar objects.



After the child named both familiar objects (i.e. ball, fish), the examiner reinforced the child's ability to name the objects by saying: *"Wow, you know all their names! Each thing here has a special name. Knowing their special names help us to tell things apart. So when someone says 'fish', I will know he is talking about this (examiner points to fish) and not this (examiner points to ball). That is why it is important to learn special names."*

Next, the examiner proceeded to direct the child to focus on the word learning task ahead (**mediation of competence – task regulation**): *"Now, Teddy (examiner points to Teddy) is going to show you some things that he likes. Some things you will know their names, but there are some things that you don't know their names. Teddy will teach you the special names of the things you don't know. You will try your best to remember them."*

7.3.1 Stage 1: Identification

At Stage 1 of the 'Teach' phase, for each target object was presented with two familiar objects (i.e. any two out of eight familiar objects that the child had named in the 'Test' phase) (Figure 7-3). To facilitate the child in identifying the target object in respond to the target word, a hierarchy of mediation prompts was provided (Figure 6-7, Chapter 6). The target word /hɑn/ is used as an example to illustrate the hierarchy of mediation prompts that was provided.

Figure 7-3. Teddy presented with a target object (referent object for /han/) and two familiar objects.

At the first level of mediation (**Level 1: contextual mediation**), the examiner provided a general elicitation question: "Teddy asks w*hich one is /han/?*" At this level, the examiner observed if the child was able to independently identify the target object in response to the target word.

If the child failed to identify the target object, the examiner proceeded to provide feedback and the next level of mediation (**Level 2: implicit context/language mediation**): *"No that is not /han/. That is a hard word. Let's find the ones you know first."* The examiner then proceeded by naming the two familiar objects while asking the child to point to them individually: *"Teddy asks which one is 'ball' (waits for the child to point), Teddy asks which one is 'spoon'?"* After both familiar objects were identified by the child, the examiner presented the target word again: *"Now, Teddy asks which one is /han/?"* Once again, the examiner observed if the child was able to identify the target object in response to the target word.

If the child still failed to identify the target object, the examiner would again provide feedback, followed by providing the final level of mediation (**Level 3: explicit context/language mediation**): *"No, this is not /han/. Let us see what we have here. This is 'ball' (examiner places Teddy behind the ball), this is /han/ (examiner places Teddy behind the unfamiliar target object*

- /han/), and this is 'spoon' (examiner places Teddy behind the spoon). Now which one is /han/?"

At this level, the target word had been explicitly matched to the unfamiliar target object. The child was expected to identify the target object correctly with the final level of mediation prompt.

For each target word, the examiner recorded the level of mediation prompts (Level 1: 1 point, Level 2: 2 points, Level 3: 3 points) required by the child to identify the target object in response to the target word. The TEACH Identification score was calculated by adding up the total number of points required to identify all four target words.

7.3.2 Stage 2: Interactive Play Activity with Teddy

After the child had correctly identified the target object, the examiner immediately proceeded to present the target word and the target object in an interactive play activity with Teddy to increase exposure to the target word (refer to Figure 7-4 below). The target word was presented nine times (indicated in superscript in the script below) in the scripted play activity. There were also three embedded opportunities for the child to imitate/use each target word during the interactive play activity. Below is an example of the target word /*han*/ presented in the interactive play activity.

Figure 7-4. Teddy and the referent object for /han/.

Example of script for /hɑn/

Once the child had identified target object in response to the target word /han/ in Stage 1, Stage 2 proceeded immediately. At the same time, the two familiar objects presented in Stage 1 would be removed immediately from the table.

Target object for /hɑn/ and teddy are on the table (Figure 7-5).

Examiner: "Yes, This is /han/[1]!" (examiner points to /han/)

Examiner: "/Han/[2] is a type of food." (examiner points to /han/)

Examiner: "What is this?" (examiner holds /han/ and shows it to the child)

First opportunity for the child to use/imitate target word. Examiner waits.

*Examiner: "Yes, /han/[3]!" (*correct response from the child*)*

OR

*Examiner: "No, /han/[3]!" (*incorrect response from the child*)*

Examiner to hold Teddy and /han/.

*Examiner: "Teddy likes to eat /han/[4]." (*examiner shows gesture of feeding Teddy /han/*)*

Examiner: "Teddy says /han/[5] tastes like chocolate!"

Examiner: "What is this?" (examiner holds /han/ and shows it to the child)

Second opportunity for the child to use/imitate target word. Examiner waits.

*Examiner: "Yes, /han/[6]!" (*correct response from the child*)*

OR

*Examiner: "No, /han/[6]!" (*incorrect response from the child*)*

Examiner encourages child to hold Teddy and /han/.

Examiner: "Now, Teddy wants you to hold /han/[7]."

*Examiner: "Show me what Teddy can do with /han/[8]." (*Encourage child to demonstrate gestural action with Teddy and /hɑn/)

Examiner: "What is this?" (examiner points to /han/)

Third opportunity for the child to use/imitate target word. Examiner waits.

*Examiner: "Yes, /han/[9]!" (*correct response from the child*)*

OR

*Examiner: "No, /han/[9]!" (*incorrect response from the child*)*

After the target word was presented nine times in the play activity with Teddy, the examiner would comment on the effort made by the child so far in learning the target word (**MLE principle: mediation of competence - praise**) by saying: *"Wow, you have learnt a special*

name for this thing (examiner to point to the target object). Try your best to remember this special name! Let's see what else Teddy has to show us!"

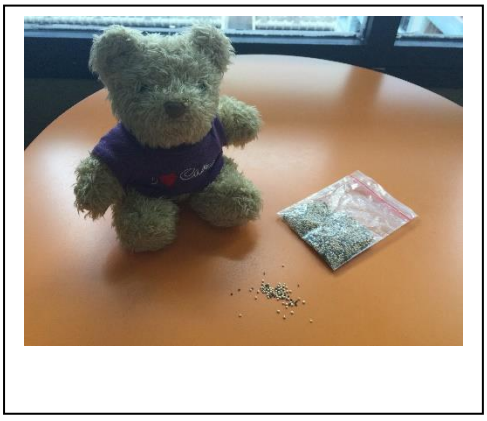
Stage 1 (Identification) and Stage 2 (scripted play activity) of the 'Teach' phase continued in cycle (refer to Figure 6-4, Chapter 6) until all four target objects in response to the target words have been identified and presented in the scripted play activity. After which, the examiner would comment on the effort (**MLE principle: mediation of competence – praise for identification stage**) made by the child in learning all four target words: *"Well done! You have listened and learned well. I can see you are trying your best to remember all the special names."*

7.3.3 Stage 3: Naming

At the third and final stage of the 'Teach' phase, the child had to name all four target objects with the target words when presented individually. The examiner introduced the naming activity by saying: *"Now, let's see if you can remember the special names of the things Teddy showed you just now*." To facilitate the child in naming the target object, a hierarchy of mediation prompts (Figure 6-8, Chapter 6) was provided. Once again, the target word /*han*/ is used to illustrate an example.

First, the examiner the target object with Teddy on the table (Figure 7-5 below), followed by asking (**Level 1: elicitation question**): *"Teddy asks what is this?"*. At this mediation level, the examiner observed if the child was able to name the target object correctly.

Figure 7-5. Teddy and the referent object for /han/.



If the child failed to name the target object with the target word after eight seconds, the examiner proceeded to provide the next level of mediation prompt (**Level 2: mediation with semantic prompt**) by describing the semantic category and/or function of the target object: *"Teddy likes to eat this. It is a of type food. What is it?"*

If the child still failed to name the target object correctly after eight seconds, the examiner proceeded to provide the next level of mediation prompt (**Level 3: mediation with phonological prompt**) by hinting the first sound of the target word: *"The name starts with /h/'. What is it?"*

If the child was still unable to name the target object correctly; the examiner would provide a final level of mediation (**Level 4: mediation with model**) by explicitly presenting the target word: *"This is /han/. What is this?"* As a verbal model had been provided by the examiner explicitly, the child was expected to at this stage to produce the target word correctly.

Regardless of the number of mediation prompts provided, once the child named the target object correctly the examiner would comment on the positive change and effort made by the child saying (**MLE principle: mediation of competence – praise**): *"Wow you have just learnt a new name. I can see that you are trying your best to remember the special name for this (holds referent object). Keep up the good work! Now let's look at another one.*"

For each target object named, the examiner recorded the level of mediation prompts (Level 1: 1 point, Level 2: 2 points, Level 3: 3 points, Level 4: 4 points) required by the child. The TEACH Naming score was calculated by adding up the total number of points required to name all four target objects with their respective target words.

After the four target objects had been named, the examiner concluded the 'Teach' phase by stating (**MLE principle: mediation of competence - praise**): *"You have worked really hard today! At first, you didn't know the names of these things (examiner to place all four target objects on table and point - Figure 7-6 below) and didn't use their special names. You were*

able to (list strategies child used, e.g., listened, looked the object, repeat to yourself) to help yourself to remember the names."

Figure 7-6. Teddy and all four target objects.

7.4 Retest Phase

The purpose of the 'Retest' phase was to evaluate the amount of change (i.e. number of target words learnt) that took place as a result of the word learning experience provided in the 'Teach' phase. In other words, the child's word learning potential. No feedback or examiner support was provided at this phase. The 'Retest' phase occurred immediately after the 'Teach' phase.

The 'Retest' phase was set up as an activity where all four target objects and four familiar objects (that child had previously named in the 'Test' phase) were placed on the table (Figure 7-7).

The examiner first assessed the child's ability to name all four target objects individually by saying: *"Teddy wants you to help keep all his things in his bag. Teddy wants you to name each thing before you place them in the bag. Try your best to use the correct name. It is all right if you can't remember the name."* As the child proceeded to name each object, the examiner recorded the number of target objects that were correctly named (i.e. RETEST Naming score).

Figure 7-7. Teddy with four familiar objects, four target objects and a bag.

After the child had named all eight objects to his/her best ability, the examiner assessed the child's ability to identify all target objects in response to the target words. The examiner proceeded by removing all eight objects from the bag and placing them on the table again before saying: "*Silly Teddy forgot what was in the bag. I have taken them out. You are going to help Teddy by pointing to the thing when I tell you the name.*" The examiner then proceeded to name each object in random order. All eight objects remained on the table regardless of whether the child was able or not able to point to the correct object when named. The examiner recorded the number of target objects that identified correctly in response to the target words (i.e. RETEST Identification score). The scoring sheet for the DA of word learning skills to record all scores is included in Appendix 4.

7.5 Chapter Summary

In this chapter, the revised version of the DA of word learning skills that was used in Phase Two of the study was presented. In the next chapter, the methodology in which the Phase Two of the study was conducted will be presented.

Chapter 8 Phase Two: Methodology

8.1 Introduction

One of the research gaps identified from the literature review (Chapter 4, Section 4.4) was that DA as an alternative assessment approach was only explored with bilingual communities who form the minorities in predominantly monolingual countries (i.e. UK, USA). Findings from these studies have consistently found that DAs, compared to static assessment measures, are more accurate in differentiating LI from language difference (i.e. typical development) in bilingual children and that it should be considered as an alternative assessment approach when assessing bilingual children (Hasson et al., 2012; Kapantzoglou et al., 2012; Peña et al., 2001). However, the feasibility and utility of DA as an assessment approach to assist clinicians/SLPs in distinguishing language difference from disorder and identifying bilingual children at risk of LI in predominantly bilingual countries have not been studied.

In Phase One of the study (as previously presented and discussed in Chapter 2 and Chapter 3), the assessment practices and challenges faced by SLPs working in a predominantly bilingual country, Singapore, were explored. The findings pointed to a need to explore alternative assessment approaches to evaluate the language skills of Singaporean bilingual children instead. It was thus fitting and a continuation from the findings from Phase One that the DA process developed in Phase Two (DA of word learning skills) was explored as an alternative assessment approach for use with Singaporean bilingual children. Specifically, whether the DA process could assist SLPs to validly evaluate the language skill of Singaporean bilingual children, and identifying those at risk of LI.

In this chapter, the methodology in which Phase Two of the study was conducted is presented. The primary aim and hypotheses of the study are first discussed in relation to the evidence required to be gathered with respect to the Validity Argument Framework (Kane, 2003; 2006).

After which, the methodology details such as participants, selection of assessment measures and data collection are further elaborated.

8.2 Purpose of the Study

To recap, the primary objective of Phase Two of the study was to evaluate whether a DA process, specifically DA of word learning skills, could validly evaluate the language skills of bilingual preschool children, and therefore identify those at risk of LI. To do so, validity evidence was gathered to support the claim under each inference to be made under the Validity Argument Framework (Kane, 2006; 2013). The evidence was then evaluated to determine the extent to which the proposed interpretation and uses of the DA of word learning skills (i.e. when a child is observed to score poorly on the DA, the child is likely to have poor language learning potential and is at risk of LI) were plausible and appropriate (Kane, 2013).

During the development of the DA process, the evidence to support the claims under *Domain Definition* and *Scoring* were gathered (i.e. presented in Chapter 5 and Chapter 6). Next, the evidence to support the claims under *Generalisation* and *Extrapolation* were gathered (Table 8-1).

Table 8-1. Claims and Assumptions for Generalisation and Extrapolation inferences under the Validity Argument Framework.

Inferences	**Claims**	**Assumptions**
Generalisation	Observed scores on the DA are consistent among bilingual children with the same language developmental profiles regardless of their prior language experience.	The scores obtained on the DA between bilingual children with and without LI are differentiated. The scores obtained on the DA by bilingual children with the same language developmental profiles are consistent regardless of their prior language experience.
Extrapolation	The construct of language learning potential as assessed on the DA predicts scores that correspond to	The scores obtained on the DA can accurately predict bilingual children's language developmental profiles.

	bilingual children's language developmental profiles.	

To gather the evidence, the following research questions were asked:

1. Is there a difference in the scores (i.e. TEACH Identification, TEACH Naming, RETEST Identification, RETEST Naming) obtained on the DA of word learning skills by typically developing Singaporean bilingual children versus those who are identified as requiring on-going speech and language therapy?
2. Can the DA of word learning skills evaluate the language skills of Singaporean bilingual children regardless of their prior language experience more accurately compared to a standardised assessment of vocabulary?
3. Can the scores obtained on the DA of word learning skills classify Singaporean bilingual children who are identified as requiring on-going speech and language therapy and those who are typically developing with high accuracy (>80.0%)?
4. Can the scores obtained on the DA of word learning skills differentiate the performance of Singaporean bilingual children with language concerns into 1) those likely at risk of LI and would require further evaluation, and 2) those with language difference (i.e. typically developing) and likely do not require further evaluation?

The hypotheses were:

1. The scores obtained on the DA of word learning skills by typically developing Singaporean bilingual children and Singaporean bilingual children who are identified as requiring on-going speech and language therapy can be differentiated. Specifically:
    a) Singaporean bilingual children who are identified as requiring on-going speech and language therapy will require more mediation prompts to identify and name

all four target words in the 'Teach' phase on the DA of word learning skills compared to typically developing Singaporean bilingual children.

b) Singaporean bilingual children who are identified as requiring on-going speech and language therapy will identify and name fewer target words in the 'Retest' phase on the DA of word learning skills compared to typically developing Singaporean bilingual children.

2. The DA of word learning skills can evaluate the language skill of Singaporean bilingual children regardless of their prior language experience more accurately compared to a standardised assessment of vocabulary.

   Specifically:

   a) On the DA of word learning skills, Singaporean bilingual children with the same developmental profiles regardless of their prior language experience will obtain similar patterns of scores. Their performance on the DA of word learning skills is less likely to be influenced by their prior language experience.

   b) On the standardised assessment of vocabulary, Singaporean bilingual children's performance will be influenced by their prior language experience. Therefore, children with the same language developmental profiles but with different prior language experience will obtain a different range of scores.

3. The set of 'Teach' (TEACH Identification, TEACH Naming) and 'Retest' (RETEST Identification, RETEST Naming) scores obtained on the DA of word learning skills can classify Singaporean bilingual children who are referred and identified to require on-going speech and language therapy and those who are typically developing with high accuracy. The classification accuracy of the DA of word learning skills will be at least 80.0% (Sensitivity: 80.0%, Specificity: 80.0%), demonstrating satisfactory classification accuracy (DiStefano & Morgan; 2011; Glascoe, 2005).

4. The performance of Singaporean bilingual children with language concerns on the DA of word learning skills can differentiate between those who are 1) likely at risk of LI and will require further evaluation, and 2) likely with language difference (i.e. typically developing) and will not require further evaluation.

8.3 Ethical Considerations

Ethics approvals (Appendix 5 and Appendix 6) were obtained from the Social and Behavioural Research Ethics Committee, Flinders University and the SingHealth Centralised Institutional Review Board, Singapore.

8.4 Study Considerations

8.4.1 Blinding

Blinding refers to the concealment of participants' group allocation from one or more individuals involved in a study (Karanicolas, Farrokhyar & Bhandari, 2010). The aim of the study was to determine whether the DA of word learning skills could validly evaluate the language skills of Singaporean bilingual children and therefore identify those at risk of LI. Therefore, it was vital that the examiner who would be evaluating children on the DA of word learning skills was blinded to the children's language developmental profiles and group allocations. This was crucial to ensure rigour and unbiased ascertainment of the outcomes of the study (Karanicolas et al., 2010). If the examiner was not blinded and had prior knowledge of the children's language developmental profiles or group allocations, the examiner could have introduced some degree of subjectivity when assessing the children on the DA of word learning skills.

In this study, the DA of word learning skills was administered by the researcher. The researcher was not involved in the recruitment and grouping of the children to ensure that she remained blinded to their language developmental profiles during the administration of the DA of word

learning skills. The children's language developmental profiles and group allocation were only revealed to the researcher after data collection was completed with all children.

8.4.2 Training of Research Assistant

For the researcher to remain blinded, the recruitment and allocation of children into groups was conducted by a Research Assistant (RA). In addition, the RA assisted with the conduct of three assessments during data collection. This was to prevent any chance of the researcher being made aware of the children's language developmental profiles prior to the administration of the DA of word learning skills.

The RA was a Singaporean English-Mandarin bilingual and a psychology graduate who had two years of working experience as an RA. She had extensive experience in the identification, recruitment and grouping of preschool children for research studies. She also had extensive experience in administering and scoring of language assessments to preschool children for data collection purposes.

Training of the RA was provided by the researcher one month prior to the identification and recruitment of potential participants for the study. The training was conducted in two parts. The first part of the training focused on familiarising the RA with the study's protocol and identification of potential participants. The following training was provided to the RA:

- Purpose of the study
- Identification of potential participants from the Development Support Program (DSP) 2016 database (refer to Section 8.5.1 for an overview of the DSP):
    - Familiarization of the inclusion and exclusion criteria (see Section 8.5.2)
    - Matching protocol (i.e. for the identification and recruitment of typically developing children – see Section 8.5.2.2)

In addition, during the actual identification and recruitment of participants, the RA could contact either the researcher's supervisors or an allocated SLP from the DSP. Both the researcher's supervisors and SLP were familiar with the protocol of the study to advise the RA on the protocol if she faced any difficulties or had any questions.

The second part of the training focused on familiarising the RA with the conduct and administration of the following assessments:

- Singapore English Action Picture Test (SEAPT; Brebner 2002)
- Primary Test of Non-Verbal Intelligence (PTONI; Ehrler & McChee, 2008)
- Clinical Evaluation of Language Fundamentals Preschool 2 United Kingdom Expressive Vocabulary Subtest (CELF P2 EV; Semel, Wiig & Secord, 2004)

During the training period, the RA conducted and administered all three assessments to five typically developing children in the presence of the researcher. This was to ensure that the RA could conduct and score the language assessments according to the respective administrative and scoring guidelines for all three assessments. Training was completed when both the researcher and RA agreed that the RA was familiar and independent with the study recruitment procedures and conduct of all three assessments. The purposes of these assessments is further elaborated in Section 8.6 in this chapter.

8.5 Participants

8.5.1 Source of Participants

The complexity of the multilingual situation in Singapore could introduce confounding language variables that would impact on the generalisability of the study results and outcomes (Creswell, 2014). To minimise the complexity of the multilingual situation in Singapore and to obtain a representative sample of the majority of the population, only Singaporean English-Mandarin speaking bilingual children (Kindergarten 1 or Kindergarten 2) of Chinese ethnicity

were recruited. In addition, these children had to be reported to have attended local preschools for at least a year to ensure that they had been exposed to the English language in formal school settings. This is because as previously elaborated in Chapter 6 (Section 6.2.2), the DA of word learning skills was designed to be conducted in the English language. Therefore, this was to ensure that there would be a fair and non-biased evaluation of the children's language learning potential on the DA of word learning skills

Children identified as requiring on-going speech and language therapy and children with language concerns were recruited from schools participating in the DSP. The DSP is a nation-wide program in Singapore that provides early identification and intervention services in preschool settings for children with mild to moderate developmental needs (Tan, Chong, Oh & Tang, 2016). For children who were screened and identified to present with language difficulties or language concerns in the DSP, intervention was provided via a speech and language therapy package or a language learning support package as elaborated below.

Speech and language therapy packages:

- Conducted by qualified SLPs.
- Children referred for speech and language therapy packages usually present with moderate to severe language difficulties that warrant immediate assessment and direct intervention.
- At the time of this study, a child who was referred for a speech and language therapy package in the DSP had received 10 weekly individual sessions with the SLP. At the end of the package, the child's initial assessment result, progress and SLP's recommendations were summarised in an end-of-intervention report.

Language learning support packages:

- Conducted by trained teachers known as Learning Support Educators (LSEds).

- Children referred for learning support packages usually present with mild language concerns that do not immediately warrant SLPs' assessment and intervention.
- At the time of this study, a child who was referred for a language learning support package in the DSP had received six or 10 weekly individual sessions with the LSEds. The sessions were either conducted individually or in small groups. At the end of the package, with the support of an SLP, if required, the child's progress and LSEd's recommendations were summarised in an end-of-intervention report.

Typically developing children were recruited from schools participating in the DSP. They were identified by matching with potential children who were identified as requiring on-going speech and language therapy (refer to Section 8.5.2.2 for matching protocol).

8.5.2 Inclusion and Exclusion Criteria

Children who received speech and language therapy packages and language learning support packages in the DSP were first identified from the DSP 2016 database. Typically developing children were identified after children from the DSP database were identified. The subsections (8.5.2.1 to 8.5.2.3) below describe the identification process, and inclusion and exclusion criteria for each group of children recruited.

Three groups of children were recruited for the study. Inclusion and exclusion criteria were established for each group (Tables 8-2 to 8-7). Inclusion and exclusion criteria checks were applied at two time points. The first sets of inclusion and exclusion criteria for the respective groups were applied when identifying potential children to be approached for recruitment (Table 8-2, Table 8-4, and Table 8-6). The inclusion and exclusion criteria were based on the information that could be obtained from preschool teachers and the DSP database. The second set of inclusion and exclusion criteria for the respective groups were applied after data collection was completed (Table 8-3, Table 8-5, and Table 8-7). The second set of inclusion

and exclusion criteria was based on information provided by parents through a parent questionnaire (refer to Section 8.6.1) and the children's performance on identification measures (refer to Section 8.6.2 and 8.6.3). The information for each child could only be gathered after parental consent was obtained for the child's participation in the study.

8.5.2.1 Speech and Language Therapy Group (SLT)

There is currently no locally available diagnostic language assessment tool that can accurately and reliably identify Singaporean bilingual children with LI (Brebner, 2010; Teoh et al., 2017). There is also a lack of defined criteria for the characteristics of LI among Singaporean bilingual children (Teoh et al., 2017). Given the limitations, the most reliable way to identify Singaporean bilingual children with LI for the study was to obtain SLPs' clinical impressions and recommendations of children who had at least completed one package of speech and language therapy in the DSP at the point of recruitment. To best identify a group of children with isolated LI, the following inclusion and exclusion criteria were applied at two time points (See Table 8-2 and Table 8-3). This group of children identified was henceforth referred to as the SLT group or SLT participants.

Table 8-2. First tier of inclusion and exclusion criteria for the SLT group (identified from DSP 2016 database).

First tier of inclusion criteria	**First tier of exclusion criteria**
• Child was referred for speech and language therapy package in the DSP and had completed at least 10 sessions of intervention in 2016. • Child of Chinese ethnicity and reported by teachers as English-Mandarin speaking. • Child was attending Kindergarten 1 or Kindergarten 2 (aged 4;0 to 7;0) at the time of recruitment. • Information gathered from the child's end-of-intervention report confirmed that the child:	Information gathered from the child's end-of-intervention report: • Child was known to have childhood developmental disorders, e.g. Autism, Global Developmental Delay. • Child was referred to speech and language therapy for articulation/speech impairment therapy only. • Child was discharged after completing one package of speech and language therapy.

✓ Had LI or was at risk of LI. ✓ Performed below average on standardised language assessments. ✓ Showed little or slow language progress after 10 sessions. ✓ Recommended to continue speech and language therapy.	

Table 8-3. Second tier of inclusion and exclusion criteria for SLT group (after data collection).

Second tier of inclusion criteria	**Second tier of exclusion criteria**
• Parents reported on the parent questionnaire that the child had attended school in Singapore for at least a year. • On the SEAPT, the child performed at or below the 20th percentile for at least one of the component (Information or Grammar raw score).	• Parents reported on the parent questionnaire that the child had a history of hearing concerns/loss. • On the PTONI, the child obtained a Non-Verbal Index score of 90 and below.

8.5.2.2 Typically Developing Group (TD)

After a potential SLT child was identified, typically developing children from the same school were identified to be approached. To minimise any potential characteristics of participants that could influence the outcomes of the study, typically developing children were identified by matching. This process was aimed to improve the rigour of the study design to minimise the differences in demographic characteristics between the groups (Creswell, 2014). The matching process is described below:

1. A potential SLT participant was first identified.

2. The RA, together with the class teacher of the SLT participant, would identify potential typically developing children from the same class or level that were matched to the potential SLT participant on:
    a. Gender
    b. Chronological age (date of birth within 6 months)
    c. Language dominance (i.e. based on teacher's observation and feedback)
3. To increase the chances of successful recruitment, up to four typically developing children were identified for each potential SLT participant.
4. Typically developing children that were identified had to meet the following inclusion and exclusion criteria at two time points (Table 8-4 and Table 8-5). This final group of children recruited was henceforth referred to as the TD group or TD participants.

Table 8-4. First tier of inclusion and exclusion criteria for TD group (identified after SLT children were identified).

First tier of inclusion criteria	**First tier exclusion criteria**
• Child identified after matching to potential SLT participant on chronological age (date of birth within 6 months), gender and language dominance. • Child of Chinese ethnicity and reported by teachers as English-Mandarin speaking. • Child was attending Kindergarten 1 or Kindergarten 2 (aged between 4;0 months to 7;0) at the time of recruitment. • Teachers reported that they had no speech, language and learning concerns with the child.	• Child had previously received DSP services.

Table 8-5. Second tier of inclusion and exclusion criteria for TD group (after data collection).

Second tier of inclusion criteria	**Second tier exclusion criteria**
• Parents reported in the parent questionnaire that the child had attended school in Singapore for at least a year.	• Parents reported in the parent questionnaire that they had specific concerns regarding their child's language development. • Parents reported in the parent questionnaire that the child had a history of hearing concerns/loss. • Parents reported in the parent questionnaire that the child received some form of early intervention or therapy. • On the PTONI, the child obtained a Non-Verbal Index score of 90 and below. • On the SEAPT, the child performed at or below the 20th percentile for at least one of the components (Information or Grammar raw score).

8.5.2.3 Language Support Group (LS)

Children with language concerns but did not receive speech and language therapy in the DSP were recruited as a separate group of participants from the SLT group and TD group. These children, at the point of recruitment, had received and completed at least one package of language learning support in the DSP. This group of children was henceforth referred to as LS group or LS participants.

One of the forms of evidence to be gathered was to determine whether the scores on the DA of word learning skills could accurately classify Singaporean bilingual children with language concerns: those likely at risk of LI who required further evaluation versus those with language difference and were not at risk of LI. This was to evaluate whether the DA of word learning skills has potential as a language screener by administering it to a group of children with language concerns (i.e. LS group). To gather the evidence, the performance of LS participants

on the DA of word learning skills was compared to the recommendations and evaluations in their end-of-intervention reports (i.e. At risk of LI: Children who were recommended for a second package of language support and further evaluation; Not at risk of LI: Children who made good progress and were discharged from DSP services).

The group of LS children were identified after applying the following inclusion and exclusion criteria at two time points (Table 8-6 and Table 8-7).

Table 8-6. First tier of inclusion and exclusion criteria for LS group.

First tier of inclusion criteria	**First tier of exclusion criteria**
• Child referred for language learning support in the DSP and had completed at least one package of intervention (six or 10 sessions) in 2016. • Child of Chinese ethnicity and reported by teachers as English-Mandarin speaking. • Child was attending Kindergarten 1 or Kindergarten 2 (aged between 4;0 months to 7;0) at the time of recruitment. • Information gathered from the child's end-of-intervention report stated that the child had: 1) Made good progress in the sessions and was discharged from DSP services. <u>OR</u> 2) Made minimum progress in the sessions and was recommended to receive additional language learning support sessions or an evaluation by a SLP.	Information gathered from the child's end-of-intervention report: • Child was known to have childhood developmental disorders e.g. Autism, Global Developmental Delay.

Table 8-7. Second tier of inclusion and exclusion criteria for LS group.

Second tier of inclusion criteria	Second tier of exclusion criteria
• Parents reported in the parent questionnaire that the child had attended school in Singapore for at least a year. • Parents reported concerns regarding their child's language development.	• Parents reported in the parent questionnaire that the child had a history of hearing concerns/loss. • Parents reported in the parent questionnaire that the child received some form of early intervention or therapy. • On the PTONI, the child obtained a Non-Verbal Index score of 90 and below.

8.5.3 Rate of Participation

A total of 117 children from 20 preschools were identified and approached to participate in the study using the first sets of inclusion and exclusion criteria (Table 8-2, Table 8-4, and Table 8-6). Parent information packs that included consent forms (Appendix 7) and letter to parents (Appendix 8) were given to the schools to be distributed to the parents of potential participants.

All 117 distributed consent forms were returned via the schools. Out of these, a total of 86 parents (73.5%) consented to their children's participation in the study while 31 parents (26.5%) declined their children's participation in the study (Table 8-8).

Table 8-8. Participation rate.

	SLT group	TD group	LS group	Total
Number of children identified and approached	24	78	15	117
Number and percentage of children whose parents provided parental consent *n* (%)	18 (75)	58 (77.3)	10 (66.7)	86 (73.5)
Number and percentage of children whose parents declined their participation in the study *n* (%)	6 (25)	20 (22.7)	5 (33.3)	31 (26.5)

8.5.4 Final Number of Participants Included in Data Analysis

The RA grouped all 86 children in accordance with the first tier of inclusion and exclusion criteria into the respective groups (SLT group, TD group, LS group). All 86 children were assessed in the study. After data collection was completed with all 86 children, the researcher no longer had to be blinded. Subsequently, the developmental profiles and group allocations of the children were revealed to the researcher. The researcher then counterchecked if all 86 children were identified and grouped correctly into their respective groups based on the first sets of inclusion and exclusion criteria (Table 8-2, Table 8-4, and Table 8-6) for the respective groups and; 2) and applied the second sets of inclusion and exclusion criteria (Table 8-3, Table 8-5, and Table 8-7).

A total number of 75 children (i.e. 18 SLT participants, 48 TD participants and 9 LS participants) were included in the final data analysis. A total of 10 TD children and 1 LS child were excluded from the final data analysis after the second tier of exclusion criteria were applied. Table 8-9 lists the reasons why these 11 children were excluded in the final analysis.

Table 8-9. Reasons for the exclusion of 11 participants in the final analysis.

Participant number	Initial group membership	Reason/s for exclusion
003	TD	• Child performed at or below 20th percentile the SEAPT • Child was of Eurasian ethnicity
009	TD	• Child previously received speech and language therapy
011	TD	• Child performed at or below 20th percentile the SEAPT • Child attended school in Singapore for less than a year at the point of recruitment
014	TD	• Child performed at or below 20th percentile the SEAPT
016	LS	• Child previously received speech and language therapy
042	TD	• Child previously attended early intervention services
064	TD	• Child performed at or below 20th percentile the SEAPT • Child obtained a Non-Verbal Index score of 90 and below on the PTONI
069	TD	• Child obtained a Non-Verbal Index score of 90 and below on the PTONI
075	TD	• Child performed at or below 20th percentile the SEAPT
081	TD	• Child performed at or below 20th percentile the SEAPT • Child obtained a Non-Verbal Index score of 90 and below on the PTONI
084	TD	• Child performed at or below 20th percentile the SEAPT

8.6 Identification Measures

The following questionnaire and assessment measures were used to screen and group children into the respective groups in the study after the parents had consented to their children's participation in the study.

8.6.1 Parents Questionnaire – Language Background Questionnaire (LBQ)

Purpose

Parents who consented to their child's participation in the study had to complete a questionnaire requesting information on their child's language background. This was to obtain information to classify participants into their respective language dominant group. The purpose of this classification was to gather evidence on whether the DA of word learning skills could evaluate the language skills of Singaporean bilingual children accurately regardless of their language experience.

The complex language environment in Singapore results in variable language experience among individuals. Each Singaporean child is exposed to a unique language learning environment depending on the number, frequency and quality of languages used in their home, school and in the community. Even within the same ethnicity group where children are likely to be exposed to at least the same two languages (i.e. Chinese children are exposed to both the English language and Mandarin language), factors such as language exposure and language use are variable. Therefore, for the purpose of research, local studies (i.e. Brebner et al., 2016; Gn, Brebner, McCormack, 2014; Pua, Lee & Rickard-Liow, 2017; Teoh et al., 2012) conducted in Singapore often group Singaporean bilingual children into subgroups based on their language dominance.

A language is considered dominant when it is preferred and used more frequently than the other language (Law & So, 2006). Language dominance can also be defined as a situation where one language is more advanced or developing faster than the other(s) (Yip & Matthews, 2006). Among bilingual children, the dominant language is usually the language they have had the most exposure to (Grosjean, 2010). Local studies have found that bilingual children who have been exposed to the same two languages but with different dominance in each, performed differently on the same language measures (Dixon, Wu & Daraghmeh, 2012; Teoh et al., 2012;

Brebner et al., 2016). SLPs may misdiagnose the presence of LI if language dominance is not considered especially if the assessment is only conducted in one language (Lim, Rickard-Liow, Lincoln, Yiong & Onslow, 2008). Therefore, it is important to determine which language is dominant when assessing bilingual children to have accurate evaluations of their actual language ability.

There is a lack of consistency on methodology in the international literature on how language dominance among bilingual children can be accurately determined (Bedore et al., 2012). In local studies (e.g. Chang, Young, Rickard-Liow & Chong, 2015; Gn et al, 2014; Pua et al., 2017; Teoh et al, 2012; Yeo, 2015), language dominance in Singaporean bilingual children has been determined using adapted versions of the Language Background Questionnaire (LBQ) originally developed by Tan (2008). Thus, the LBQ was selected as an appropriate questionnaire to determine the language dominance of the participants for the current study.

For the current study, the language background information gathered from the LBQ was compared with the teacher's verbal feedback to determine each child's dominant language. Using multi-sourced information reported from both parents and teachers provided a comprehensive language background check, as compared to using only a single informant source (Sheng, Lu & Gollan, 2014; Teoh et al., 2012). Participants in the study were classified either English Dominant (ED) or Mandarin Dominant (MD).

Administration

The LBQ, together with parental consent forms and information letters were distributed to potential participants. Parents were informed in the consent form to complete and return the LBQ if they were willing to consent to their child's participation in the study.

The following information was gathered from the LBQ (Appendix 9):

- Detailed information on the child's language background:

    - The number of languages the child was exposed to
    - Age at which the child was first exposed to each language
    - Languages used by the child's main caregivers and the estimated percentage use of each language with the child
    - Parent's perception of their child's understanding and speaking proficiency in each language
- Demographic information such as:
    - Housing status
    - Parents' highest educational attainment
- Other pertinent information such as:
    - Number of years the child had been attending local preschool
    - Parental concerns, if any, over child's hearing, speech and language development

8.6.2 Singapore English Action Picture Test (SEAPT)

Purpose

Children whose parents consented to their participation in the study were screened on a language screening measure. This was to ensure that 1) potential TD participants had average language skills and were not at risk of LI (i.e. performed above the 20th percentile on for both Grammar and Information components on the SEAPT), and 2) potential SLT participants had below average language skills and were identified at risk of LI (i.e. performed at or below the 20th percentile for either or both Grammar and Information components on the SEAPT).

The SEAPT was selected as an appropriate language screener for this purpose as it is the only locally developed language screening assessment tool that has been validated for use with Singaporean English-Mandarin preschool children based on their dominant language (ED or MD) (Brebner, 2002).

Administration of the SEAPT

All children were assessed using the SEAPT to provide a sample of their expressive language skills in English at the sentence level. The children were shown 10 pictures and were asked to describe each picture. The SEAPT allowed a range of target vocabulary and grammatical structures in English to be elicited from the participants.

The SEAPT was administered and scored as per instructions outlined in the manual. Scores were tabulated as 'Information raw scores' and 'Grammar raw scores'. Each child's scores were compared to the norms available for their age and dominant language (i.e. ED or MD). Children who scored at or below the 20th percentile for their age and dominant language were at risk of LI.

For potential TD participants to be included in the final data analysis, they had to score above the 20th percentile for their age and dominant language for both scoring components. For potential SLT participants to be included in the final data analysis, they had to score below the 20th percentile for their age and dominant language for at least one of the scoring components.

8.6.3 Primary Test of Non-verbal Intelligence (PTONI)

Purpose

Children whose parents consented to their participation in the study were screened on a non-verbal screening measure. A non-verbal screen was required to ensure all potential participants had non-verbal intelligence in the average range and were not at risk of other comorbidities (i.e. Global Developmental Delay, cognitive delay).

The PTONI was selected as it has been used in language studies (e.g. Carson, Gilon & Boustead, 2011; Messier & Wood, 2015) where quick and reliable screens of the non-verbal intelligence skills of the participants were required. The PTONI can be administered with minimal oral instructions (i.e. point to the one that is different) and the children only had to respond by

pointing. This provided a fair evaluation of non-verbal intelligence even for children whose dominant language was not English but had at least one year of exposure to the English language in school.

Administration

All children were assessed using the PTONI to provide a measure of their non-verbal cognitive skills. The PTONI was administered and scored as per instructions outlined in the manual. The children's scores were tabulated as raw scores before being converted and compared to Non-Verbal Index scores available for their age. Children with low Non-Verbal Index scores (i.e. below 90) on the PTONI were identified as 'below average' and potentially at risk of cognitive delay. For children to be included in the final data analysis, they had to obtain Non-Verbal Index scores above 90.

8.7 Study Measures

The following assessment measures were used to explore the primary aim of Phase Two of the study which was to determine whether the DA of word learning skills as an alternative assessment approach could validly identify bilingual children at risk of LI in comparison to a commercially-available standardised assessment of vocabulary.

8.7.1 Dynamic Assessment of Word Learning Skills (DA of word learning skills)

Purpose

The DA of word learning skills was developed through a research process that involved an extensive literature review and a pilot study (see Chapter 4, 5, 6). The purpose was to compare the performance of typically developing Singaporean bilingual children on the DA of word learning skills and Singaporean bilingual children who were identified as requiring on-going speech and language therapy. This was to determine whether the DA of word learning as compared to a standardised assessment of vocabulary could validly evaluate the language skills

of Singaporean bilingual children, and therefore identify those at risk of LI. The results gathered would be used as evidence to support the claims under a chain of inferences under Kane's Validity Argument Framework (2006; 2013).

Administration

All children were assessed on the DA of word learning skills as outlined in Chapter 7.

8.7.2 Articulation Screener

Purpose

The purpose of the articulation screener was to provide a quick screen of the children's speech skills before they were assessed on the DA of word learning skills. This was to determine if the children presented with any consistent speech errors that might impact their production of the target words on the DA of word learning skills. If consistent speech errors (i.e. phonological processes, phonemic errors) were observed on the articulation screener and DA of word learning skills, their responses would be considered correct.

A 5 minute articulation screener was developed (Appendix 10) by the researcher for the purpose of the study. The articulation screener was adapted and modified from two commercially-available articulation screeners (Preschool Language Scales Edition 4 United Kingdom Articulation Screener, 2002; The Quick Screener, 1996)

Administration

All children were assessed on an articulation screener. The articulation screener consisted of 11 familiar English words that the children were asked to repeat after the examiner's verbal models. The 11 words were derived from the consonants and vowels of the four target words to be learnt on DA of word learning skills. For example, for the target word /hɑn/, the English words '**h**igh', 'f**a**rm' and 'yaw**n**' were used. Speech errors, if any, were recorded qualitatively.

8.7.3 Clinical Evaluation of Language Fundamentals Preschool 2 Expressive Vocabulary Subtest (CELF P2 EV)

Purpose

One of the forms of evidence to be gathered was to compare the children's performance on the DA of word learning skills to their performance on a standardised test of vocabulary. This allowed the evaluation of whether the DA of word learning skills could better capture the language developmental profiles of Singaporean bilingual children as compared to a commercially available standardised assessment of vocabulary that was not normed or created for use with the Singaporean population.

The CELF P2 was selected as the initial survey (Phase One of the study) identified that it was one of the most common standardised assessments used by SLPs in Singapore to assess the language skills of Singaporean children (Teoh et al., 2017).

Administration

All children were assessed on the CELF P2 EV subtest. The subtest required the children to name 20 pictures to his or her best ability. The CELF P2 EV subtest was administered and scored as per the instructions outlined in the test manual. The children's scores were tabulated as raw scores before being converted to scaled scores available for their age.

8.8 Data collection

8.8.1 General procedure

Children whose parents provided consent to their child's participation in the study were assessed. Each child participated in two sessions of assessment of 30-45 minutes each. The sessions for each child were conducted at least three days apart to prevent participant fatigue. All children were assessed in a quiet room in their individual schools on regular school days.

Session 1 was conducted by the RA. Session 2 was conducted by the researcher who is a qualified SLP.

Before each session, verbal assent was obtained in the room in the presence of a teaching staff. All children provided their verbal consent before each session. The series and sequence of activities and assessments that each session had to complete for each session are listed below. All sessions were audio-recorded to ensure procedural integrity.

Session 1:

1. Play activity (5 minutes)

   To establish rapport between the examiner and the child before any evaluation.

2. Introduction to the activity chart (1 minute)

   To introduce to the child the number of activities he or she had to complete in the session.

3. Evaluation of the child on the CELF P2 EV subtest (5-10 minutes)
4. Evaluation of the child on the SEAPT (10 minutes)
5. Evaluation of the child on the PTONI (5-15 minutes)
6. Play activity and reward (3 minutes)

Session 2:

1. Play activity (5 minutes)

   To establish rapport between the examiner and the child before any evaluation.

2. Introduction to the activity chart (1 minute)

   To introduce to the child the number of activities he or she had to complete in the session.

3. Evaluation of the child on the articulation screener (5 minutes)
4. Evaluation of the child on the DA of word learning skills (20 – 30 minutes)

5. Play activity and reward (3 minutes)

8.8.2 Inter-rater reliability

All sessions were audio-recorded to ensure that the assessments were administered according to schedule and to ensure procedural integrity. The researcher reviewed 15 sessions (out of 75; 20%) conducted by the RA.

As part of evaluating the procedural integrity, assessments (SEAPT, CELFP2 EV subtest) that evaluated expressive outputs from the participants were also reviewed for inter-rater reliability. The DA of word learning skills, PTONI could not be reviewed for inter-rater reliability as both assessments recorded receptive responses from (pointing) participants that were not captured on audio-recordings.

The researcher reviewed the audio recordings and re-scored the SEAPT and CELFP2 EV subtest performance of 15 children. Pearson correlation tests were used to calculate inter-rater reliability. High inter-rater coefficients were obtained for the SEAPT ($r = 0.963$, $p < 0.001$) and CELF P2 EV subtest ($r = 1.00$, $p < 0.001$).

8.9 Chapter Summary

The primary aim of Phase Two of the study was to evaluate whether a DA of word learning skills can validly evaluate the language skills of bilingual children and therefore identify those at risk of LI as compared to a standardised assessment of vocabulary. To do so, the performance of typically developing English-Mandarin bilingual preschool children and English-Mandarin bilingual preschool children who were identified for on-going speech-language therapy on the DA of word learning skills and a standardised assessment of vocabulary (CELF P2 EV) were explored. The results are presented in the next chapter.

Chapter 9 Phase Two: Results

9.1 Introduction

This chapter presents the results of Phase Two of the study. The results from SLT and TD participants are reported in the first section (Section 9.2). The demographic characteristics of SLT and TD participants are first presented, between-group differences are highlighted, if any. The performance of SLT and TD participants on the DA of word learning skills, CELF P2 EV subtest, SEAPT and PTONI are subsequently statistically analysed and compared. Last but not least, the classification accuracies of the DA of word learning skills and CELF P2 EV subtest are compared. The results from the LS group are then presented and reported as an additional of participants to further support the classification accuracy of the DA of word learning skills.

9.2 Demographics of the Speech Language Therapy Group and Typically Developing Group

Demographic information was obtained from the LBQ completed by parents of the participants and the DSP database. The following demographic details between SLT and TD group were compared and analysed:

- Chronological age
- Gender
- Language dominance
- Number of years in school
- Social-economic status (based on housing status)
- Mother's highest education qualification

9.2.1 Analysis

Descriptive statistics were reported for each group separately. Due to small and unequal sample sizes of children within each group, bootstrapping (1000 samples at 95% confidence interval)

was applied when comparing group differences (Haukoos & Lewis, 2005). The between-group differences were assessed using One Way ANOVA (i.e. when dependent variables are continuous) or Likelihood Ratio Chi-Square (i.e. when variables are categorical). All analyses were conducted using SPSS for Windows (Version 23).

9.2.2 Chronological Age

The difference in ages between groups was not statistically significant ($F_{(1,64)} = 1.722$, $p = 0.194$). This was expected as TD participants were matched to SLT participants on chronological age (date of birth within 6 months). Table 9-1 below shows the mean age of participants in the SLT and TD group respectively.

Table 9-1. Mean age for SLT group and TD group.

	SLT group	**TD group**	***p***
Mean Age (years; months)	5;10	6;0	0.194

9.2.3 Gender

There was no significant difference between the SLT and TD group in the number of males and females ($X^2(1) = 3.125$, $p = 0.077$) (Table 9-2). This was expected as TD participants were matched to SLT participants on gender.

There were more males than females in both groups. The high number of males in the SLT group is consistent with the higher prevalence of LI in boys in comparison to girls (i.e. Tomblin et al., 1997).

Table 9-2. Numbers of males and females in SLT and TD group.

	SLT group	**TD group**	***p***
Gender (numbers of males: number of females)	16:2	33:15	0.077

9.2.4 Language Dominance

Language dominance was determined by comparing teacher's verbal/written feedback on child's dominant language with the parent's overall report of the child's language dominance on the LBQ. As mentioned previously (Chapter 8, Section 8.6.1), using reported language information provided from both parents and teachers provided a comprehensive language background check. Out of 66 pairs of report/feedback received, 46 pairs (69.7%) of parents' and teachers' feedback matched. For the remaining 20 pairs (30.3%) that were unmatched, parental responses on the LBQ for each individual child were further analysed. For all remaining cases, the parental report on their child's dominant language was chosen as the child's dominant language as more detailed information on language use, and exposure were collected via the LBQ.

There was no significant difference between the SLT and TD group in the number of MD and ED participants ($X^2(1) = 0.578$, $p = 0.447$). This was expected as TD participants were matched to SLT participants on language dominance (i.e. using teacher's verbal feedback on child's dominant language first to identify potential TD participants with likely the same dominant language). Table 9-3 below shows that there were an equal number of ED and MD participants in the SLT group. There were slightly more MD participants than ED participants in the TD group.

The overall division (i.e. slightly more MD participants compared to ED participants) of ED and MD participants in both groups was also comparable to Census data (Mandarin: 46.1%; English: 37.4%; Others: 16.5%) on languages used most commonly at home among the Singaporean Chinese ethnic population (Singapore Statistics Board, 2014).

Table 9-3. Number of ED children and MD children in each group.

	SLT group	TD group	*p*
Dominant language (numbers of ED children: number of MD children)	9:9	19:29	0.447

9.2.5 Number of Years Attending Local Preschools

The difference between groups in the number of years participants had been attending local preschools was statistically significant ($F_{(1,61)}$= 4.081, $p < 0.05$). TD participants had been attending local preschools longer compared to SLT participants. Nonetheless, on average, both TD and SLT participants had been attending local preschools between 3 to 4 years (Table 9-4).

Table 9-4. Means numbers of years attending local preschool in each group.

	SLT group	TD group	*p*
Mean number of years in school	3.17	3.78	<0.05

9.2.6 Social-Economic Status (Housing type)

Social-economic status was assessed using housing type. In Singapore, this is often used as a measure of socio-economic status in local studies (e.g. Low et al., 2016; Sababayagam, Shankar, Wong, Saw & Foster, 2008) as housing type and socio-economic status correlate for this population (Singapore Statistics Board, 2012).

In Singapore, approximately 82.0% of the population live in public housing known as Housing Development Board (HDB) flats that vary in size from one-room to five-room or more (Singapore Statistics Board, 2014). The rest of the population (18.0%) live in private housing such as private apartments and landed properties (i.e. houses). As a general trend, the low income to higher middle-income households (i.e. total monthly income of up to SGD$12,000) live in HDB flats while the high-income households (i.e. total monthly income more than SGD $10,000) live in private housing. Furthermore, households with lower income generally live

in smaller HDB flats (i.e. two-room and below) as these flats are highly subsidised by the government and can only be purchased by households with a total monthly income of not more than SGD $6,000 a month.

There was a significant difference between the TD and SLT group in housing type ($X^2(4) = 13.450$, $p < 0.05$). Table 9-5 below shows that the overall majority (88.9%) of the SLT participants lived in HDB four-room or below housing dwellings whereas the majority of the TD participants (81.3%) lived in HDB four- room or above housing dwellings. The housing types of TD participants were comparable to that of Census data (Singapore Statistics Board, 2014) whereas the housing types of SLT participants were not comparable to Census data. This observation was consistent with the literature that children from lower socio-economic status backgrounds are more at risk of LI (Tomblin et al., 1997; Kelly, 2014).

Table 9-5. Housing type status in each group (number and percentages).

	Housing type					
	HDB 1/2 room	**HDB 3 room**	**HDB 4 room**	**HDB 5 room/executive**	**Private housing**	**Not provided**
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**
SLT group	1 (5.6)	6 (33.3)	9 (50.0)	1 (5.6)	0 (0)	1 (5.6)
TD group	1 (2.1)	8 (18.8)	15 (31.3)	16 (33.3)	8 (16.7)	0 (0)
Census data (2014)	- (5)	- (19.0)	- (33.0)	- (25.0)	- (18.0)	Not applicable

9.2.7 Mother's Highest Educational Qualification

There was a significant difference in mother's highest educational level between the SLT and TD group ($X^2(5) = 15.013$, $p < 0.01$). The majority of the mothers of SLT participants' (61.1%) had attained a Secondary and below educational qualification, whereas the majority of the mothers (50.0%) of TD participants attained a University qualification. Moreover, the percentage division seen in mother's highest educational qualification of TD participants was

comparable to that of Census data (Singapore Statistics Board, 2010) whereas the percentage division mother's highest educational qualification of SLT participants was not comparable to Census data (Table 9-6). Once again, this observation was consistent with the literature that children whose mothers have lower educational attainment are at higher risk of LI (Reilly et al., 2014).

Table 9-6. Reported mother's highest educational level (numbers, percentages) for each group.

	Mother's highest educational qualification					
	Primary	**Secondary**	**Post-Secondary (non-tertiary)**	**Diploma & Professional qualification**	**University**	**Not Provided**
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**
SLT group	3 (16.7)	8 (44.4)	0 (0.0)	3 (16.7)	3 (16.7)	1 (5.6)
TD group	2 (4.2)	6 (12.5)	3 (6.3)	13 (27.1)	24 (50.0)	0 (0.0)
***Census data (2010)**	(6.1-6.4)	(12.4-19.3)	(10.7-12.0)	(24.2-17.8)	(34.6-46.5)	Not applicable

* Reported as highest education qualification by age group (25 to 34 – 35 to 44).

9.2.8 Section Summary

SLT participants were matched to TD participants on chronological age, gender and language dominance (i.e. ED or MD). As expected, there was no significant difference in these characteristics between the two groups of participants. The methodology of matching had minimise the occurrence these three confounding variables that may influence their performance on the DA of word learning skills.

There was a significant difference in the socio-economic status (housing type) and mother's highest education qualifications between two groups. This was anticipated as children with LI are likely to come from more disadvantaged backgrounds (i.e. lower SES, lower mother's education qualification) (Kelly, 2014; Reilly et al., 2014; Tomblin et al., 1997). There was also

a significant difference in the number of years in attending local preschools between SLT and TD participants.

9.3 Assessments

9.3.1 Analysis

Each assessment measure was scored and compared between groups. For the DA of word learning skills, performance at the 'Teach' phase (i.e. level of mediation prompts provided to identify and name all four target words) and 'Retest' phase (i.e. number of target words identified and named independently) were also scored and compared between groups. Due to unequal sample sizes, bootstrapping (1000 samples at 95% confidence interval) was applied when comparing group differences (Haukoos et al., 2005). Between-group differences on the DA of word learning skills were evaluated using a One-Way MANCOVA while between-group differences on the other assessments were evaluated using a One-Way ANCOVA. All analyses were conducted using SPSS for Windows (Version 23). Effect sizes were reported using partial eta squared (partial η2) for more focused comparisons between groups, dependent variables and covariates (Field, 2013).

Although matching of participants was attempted, there was an unequal number of children in each group as 1) more TD children were identified to be approached to increase the chances of successful recruitment, and 2) the rate of participation in each group could not be controlled. Thus, it was crucial that the analyses were controlled for potential confounding factors. Ideally, analyses should control for all six possible demographic factors (i.e. chronological age, gender, language dominance, number of years attending local preschools, SES, mother's highest educational qualification). However, due to the small sample size (n = 66), the number of covariates to control for had to be limited. This is because a degree of freedom is lost for every additional covariate added in the formula which in turns reduces the statistical efficiency (statistical power) of ANCOVA and MANCOVA procedures (Field, 2013). Statistical experts

(e.g. Field, 2013; Miller & Chapman, 2001) have suggested that the covariates that are to be controlled for must be independent of the grouping variable (i.e. TD or SLT). Gender bias (more boys than girls), low SES levels and low mother's highest educational qualification levels are demographic characteristics of children with LI (Kelly, 2014; Reilly et al., 2014; Tomblin et al., 1997). In other words, gender, SES and mother's highest education qualification are closely correlated to the risk of LI. Therefore these characteristics should not be controlled when evaluating the differences in groups' performance as it would impact the outcomes of the study, leading to significant cost to any potential generalization of the outcomes of this study. The covariates chronological age, language dominance and the number of years attending local preschools, however, were independent of the grouping variable (i.e. SLT or TD) and are controlled for when evaluating the differences in groups' performance.

9.3.2 Articulation Screener (Study Measure)

Out of the 66 participants, 61 (92.1%) were able to produce all 11 English words on the screener accurately. This indicated that the majority of the participants did not present with speech difficulties that may impact on their production of the target words on the DA of word learning skills. The other five participants (7.9%) were observed to have either one to two speech errors (i.e. phonological or phonemic). These five participants' speech errors were recorded and were not penalised on the DA of word learning skills if their production errors of the target words were consistent with their speech errors observed on the articulation screener.

9.3.3 Dynamic Assessment of Word Learning Skills (Study Measure)

9.3.3.1 Speech and Language Therapy Group versus Typically Developing Group

The adjusted means and standard errors for word identification and naming for each group for the 'Teach' and 'Retest' phases are presented in Table 9-7. As anticipated, SLT participants required more mediation prompts than TD participants in both identifying and naming all four

target words in the 'Teach' phase. Thus the means for the SLT group were higher than the TD group in the 'Teach' phase. The between-group difference (after adjustment for chronological age, language dominance and number of years attending local preschools) was significant for TEACH Identification score ($F_{(1, 58)}$ = 8.240, p <0.05, partial η2 = .124) and TEACH Naming score ($F_{(1, 58)}$ = 4.604, p <0.05, partial η2 = .067).

Cohen's (1969, pp. 278-280) criteria of small (.0099), medium (.0588) and large effect sizes (.1379) to interpret partial eta square values (Richardson, 2011) were used to interpret the effect sizes. The difference between SLT and TD group in TEACH Identification score represented a medium to large effect size (partial η2 = .124) whereas the difference between SLT and TD group in TEACH Naming score (partial η2 = .067) represented a medium effect size.

As anticipated, SLT participants identified and named fewer target words in the 'Retest' phase compared to their TD counterparts. Thus the means for the SLT group were lower than the TD group in the 'Retest' phase. The between-group differences after adjustment for chronological age, language dominance and number of years attending local preschools) were significant for both RETEST Identification score ($F_{(1, 58)}$ = 102.163, p < 0.001, partial η2 =.638) and RETEST Naming score (Retest) ($F_{(1, 58)}$ = 30.472, p < 0.001, partial η2 = .344).

Cohen's (1969, pp. 278-280) criteria of small (.0099), medium (.0588) and large effect sizes (.1379) to interpret partial eta square values (Richardson, 2011) were used to interpret the effect sizes. The difference between SLT and TD group in RETEST Identification score represented a large effect size (partial η2 = .638). Similarly, the difference between SLT and TD group in RETEST Naming score also represented a large effect size (partial η2 = .344).

The covariates (chronological age, language dominance and the number of years attending local preschools) were all not significantly correlated to the participants' performance on the DA of word learning skills.

Table 9-7. Adjusted Means (M) and Standard Errors (SE) on the DA of word learning skills for SLT group and TD group.*

	SLT group (*n* = 19)	**TD group (*n* = 45[#])**		
	M (SE)	**M (SE)**	***p***	η2
TEACH Identification score	4.741 (0.186)	4.104 (0.115)	<0.05	.124
TEACH Naming score	14.219 (0.737)	12.601 (0.458)	<0.05	.067
RETEST Identification score	1.875 (0.165)	3.872 (0.103)	<0.001	.638
RETEST Naming score	0.355 (0.265)	2.102 (0.165)	<0.001	.344

*Adjusted scores after controlling for chronological age, language dominance and number of years attending local preschools.

[#]There were 3 participants whose parents did not provide the number of years their children had been attending local preschools.

9.3.3.2 Typically Developing English Dominant participants versus Typically Developing Mandarin Dominant participants

The adjusted means and standard errors for word identification and naming for TD ED and TD MD participants for the 'Teach' and 'Retest' phases are presented in Table 9-8. The between-group differences between TD ED participants and TD MD participants in all measurements of the DA of word learning skills were statistically non-significant (i.e. TEACH Identification, p = 0.574; TEACH Naming, p = 0.801; RETEST Identification, p= 0.682; RETEST Naming, p =.840).

Table 9-8. Adjusted Means(M) and Standard Errors (SE) on the DA of word learning skills for TD ED and TD MD participants.*

	TD English Dominant (*n* = 19)	**TD Mandarin Dominant (*n* = 26[#])**	
	M (SE)	**M (SE)**	***p***
TEACH Identification score	4.052 (0.085)	4.116 (0.073)	0.574
TEACH Naming score	12.660 (0.766)	12.402 (0.653)	0.801
RETEST Identification score	3.904 (0.120)	3.839 (0.102)	0.682
RETEST Naming score	2.067 (0.287)	2.144 (0.244)	0.840

*Adjusted scores after controlling for chronological age, language dominance and number of years attending local preschools.

[#]There were 3 participants whose parents did not provide the number of years their children had been attending local preschools.

9.3.4 Clinical Evaluation of Language Fundamentals Preschool 2 Expressive Vocabulary Subtest (Study Measure)

9.3.4.1 Speech and Language Therapy Group versus Typically Developing Group

The adjusted means and standard errors for CELF P2 EV scaled scores for each group are presented in Table 9-10. As anticipated, SLT participants scored more poorly than their TD counterparts. The between-group difference (after adjustment for chronological age, language dominance and number of years attending local preschools) was significant ($F_{(1, 58)}$ = 42.599, p < 0.001, partial η2 = .423). However, unlike for the DA of word learning skills, the covariate language dominance was significantly co-related to the participants' performance on the CELF P2 EV subtest ($F_{(1, 58)}$ = 28.695, p < 0.001, r =0.58). The large effect (r >0.5; Field, 2013) of the covariate language dominance means that the children's dominant language had influenced their performance on the CELF P2 EV subtest.

Table 9-9. Adjusted Means (M) and Standard Errors (SE) on the CELF P2 EV subtest for SLT group and TD group.*

	SLT group (*n* = 19)	**TD group (*n* = 45[#])**		
	M (SE)	**M (SE)**	***p***	η2
CELF P2 EV scaled score	5.175 (0.478)	8.908 (0.297)	<0.001	.423

*Adjusted scores after controlling for chronological age, language dominance and number of years attending local preschools.

[#]There were 3 participants whose parents did not provide the number of years their children had been attending local preschools.

9.3.4.2 Typically Developing English Dominant Participants versus Typically Developing Mandarin Dominant Participants

TD ED participants on average achieved higher scaled scores on the CELF P2 EV compared to TD MD participants (Table 9-11). The between-group difference (after adjustment for chronological age, and number of years attending local preschools) was statistically significant ($F_{(1, 41)}$ = 21.571, p < 0.001, partial η2 = .345).

Table 9-10. Adjusted Means (M) and Standard Errors (SE) on the CELF P2 EV subtest for TD ED and TD MD participants.*

	TD English dominant (*n* = 19)	**TD Mandarin dominant (*n* = 26[#])**		
	M (SE)	**M (SE)**	***p***	η2
CELF P2 EV scaled score	10.518 (0.484)	7.544 (0.413)	<0.001	.345

*Adjusted scores after controlling for chronological age, language dominance and number of years attending local preschools.

[#]There were 3 participants whose parents did not provide the number of years their children had been attending local preschools.

9.3.4.3 Speech and Language Therapy English Dominant Participants versus Typically Developing English Dominant Participants

SLT ED participants on average achieved higher scaled scores on the CELF P2 EV subtest compared to SLT MD participants (Table 9-12). The between-group difference (after adjustment for chronological age, and number of years attending local preschools) was significant ($F_{(1, 14)} = 8.014$, $p < 0.05$, partial η2 = .364).

Table 9-11. Adjusted Means (M) and Standard Errors (SE) on the CELF P2 EV subtest for SLT ED and SLT MD participants.*

	SLT English dominant (*n* = 9)	**SLT Mandarin dominant (*n* = 9)**		
	M (SE)	**M (SE)**	***p***	η2
CELF P2 EV scaled score	6.601 (0.571)	4.288 (0.571)	<0.05	.364

*Adjusted scores after controlling for chronological age and number of years attending local preschools.

9.3.4.4 Speech and Language Therapy English Dominant Participants versus Speech and Language Therapy Mandarin Dominant Participants

The adjusted means and standard errors for word identification and naming for SLT ED and SLT MD participants for the 'Teach' and 'Retest' phases are presented in Table 9-9. The between-group difference between SLT ED participants and SLT MD participants in all measurements of the DA of word learning skills were statistically non-significant (i.e. TEACH

Identification, p = 0.648; TEACH Naming, p = 0.526; RETEST Identification, p= 0.816; RETEST Naming, p =.524).

Table 9-12. Adjusted Means and Standard Errors on the DA of word learning skills for SLT ED and SLT MD participants.*

	SLT English Dominant (n = 9)	**SLT Mandarin Dominant (n = 9)**	
	M (SE)	**M (SE)**	***p***
TEACH Identification score	4.916 (0.415)	4.639 (0.415)	0.684
TEACH Naming score	14.823 (0.814)	14.066 (0.814)	0.526
RETEST Identification score	1.836 (0.311)	1.942 (0.311)	0.816
RETEST Naming score	0.433 (0.214)	0.234 (0.214)	0.524

*Adjusted scores after controlling for chronological age and number of years attending local preschools.

9.3.5 Primary Test of Non-Verbal Intelligence (Identification Measure)

9.3.5.1 Speech and Language Therapy Group versus Typically Developing Group

After adjustment for chronological age, language dominance and the number of years attending local preschools, there was a significant statistical difference between the non-verbal performance of SLT and TD participants ($F_{(1, 58)}$ = 9.983, p < 0.05, partial η2 = .147). Although both groups of participants performed within the average range (Non-Verbal Index Score >90), TD participants obtained a higher Non-Verbal Index Score compared to their SLT counterparts (Table 9-13).

Table 9-13. Adjusted Means (M) and Standard Errors (SE) on the PTONI for SLT group and TD group.*

	SLT group (n = 18)	**TD group (n = 45[#])**		
	M (SE)	**M (SE)**	***p***	η2
PTONI Non-Verbal Index score	114.16 (3.14)	126.03 (1.95)	<0.05	.147

*Adjusted scores after controlling for chronological age, language dominance and number of years attending local preschools.

[#]There were 3 participants whose parents did not provide the number of years their children had been attending local preschools.

9.3.6 Singapore English Action Picture Test (Identification Measure)

9.3.6.1 Speech and Language Therapy Group versus Typically Developing Group

After adjustment for chronological age, language dominance and the number of years attending local preschools, there was a statistically significant difference in SEAPT Information raw scores between SLT and TD participants ($F_{(1, 58)} = 31.018$, $p < 0.001$, partial η2 = .348). There was also a statistically significant difference in SEAPT Grammar raw scores between SLT and TD participants ($F_{(1, 58)} = 9.038$, $p < 0.05$, partial η2 = .135). TD participants on average obtained higher Information raw scores and Grammar raw scores compared to their SLT counterparts (Table 9-14). This result was anticipated as the SEAPT is a local expressive language tool used for identifying Singaporean English-Mandarin preschool children who are at risk of LI for further assessment. SLT participants had been previously identified to continue on-going speech and language therapy and therefore were expected as a group to perform below their typically developing peers on the SEAPT.

Table 9-14. Adjusted Means(M) and Standard Errors (SE) on the SEAPT for SLT group and TD group.*

	SLT group (*n* = 18)	**TD group (*n* = 45[#])**		
	M (SE)	**M (SE)**	***p***	η2
SEAPT Information raw score	54.78 (1.53)	64.96 (0.95)	<0.001	.348
SEAPT Grammar raw score	34.71 (2.26)	42.86 (1.41)	<0.05	.135

*Adjusted scores after controlling for chronological age, language dominance and number of years attending local preschools.

[#]There were 3 participants whose parents did not provide the number of years their children had been attending local preschools.

9.4 Classification accuracy

9.4.1 Analysis

Descriptive statistics were reported when describing classification accuracy. Logistic regression was conducted to assess whether scores on the DA of word learning skills classified

participants accurately into their respective groups (SLT, TD). Due to unequal sample sizes, bootstrapping (1000 samples at 95% confidence interval) was applied when running the model. All analyses were conducted using SPSS for Windows (Version 23).

9.4.2 Dynamic Assessment of Word Learning Skills

Binomial logistic regression was performed to evaluate how accurately the participants' performance on the DA of word learning skills could predict their group membership (SLT or TD). The logistic regression model was statistically significant, $\chi 2(4) = 55.244$, $p < .001$. The model explained 82.1% (Nagelkerke R^2) of the variance in group membership and correctly (predicted) classified 93.9% of the participants (Table 9-15). Sensitivity was 88.9% with 16 out of 18 SLT participants classified as SLT participants.

Table 9-15. Classification accuracy of the DA of word learning skills.

	Predicted		
	TD (*n*)	**SLT (*n*)**	**Percentage correct**
TD (*n*)	46	2	Specificity: 95.8
SLT (*n*)	2	16	Sensitivity: 88.9

9.4.3 Clinical Evaluation of Language Fundamentals Preschool 2 Expressive Vocabulary Subtest (Study Measure)

The number of TD and SLT participants that performed below average (i.e. at risk of LI) on the CELF P2 EV subtest was evaluated (Table 9-16 and Table 9-17). According to the CELF P2 manual, a scaled score of 6 and below indicates that the child has performed below average and is at risk of mild to severe LI.

A total of 11 TD participants (22.9%) performed below average on the CELF P2 EV subtest. On further analysis, all 11 participants were MD. A total of 12 SLT participants (66.7%) performed below average on the CELF P2 EV subtest, of which eight were MD while the other four children were ED.

The overall classification accuracy was at 74.2% (Table 9-18). Sensitivity was 66.7% with 12 out 18 SLT participants classified as SLT participants (scaled score of 6 and below). Specificity was 77.1% with 37 out of 48 TD participants classified as TD participants (scaled score of 7 and above).

Table 9-16. Performance of the TD participants on the CELF P2 EV subtest.

CELF P2 EV scaled scores	Number of TD participants	Percentage
#4	4	8.3
#5	2	4.2
#6	5	10.4
7	6	12.5
8	4	8.3
9	7	14.6
10	7	14.6
11	5	10.4
12	4	8.3
13	4	8.3

#At risk of LI (below 16th percentile based on UK normative data).

Table 9-17. Performance of the SLT participants on the CELF P2 EV subtest.

CELF P2 EV scaled scores	Number of SLT participants	Percentage
#2	1	5.6
#3	3	16.7
#4	2	11.1
#5	5	27.8
#6	1	5.6
7	3	16.7
8	1	5.6
9	1	5.6
10	1	5.6

#At risk of LI (below 16th percentile based on UK normative data)

Table 9-18. Classification accuracy of the CELF P2 EV subtest.

	Actual (based on scaled scores)		
	TD (*n*)	**SLT (*n*)**	**Percentage correct**
TD (*n*)	37	11	Specificity: 77.1
SLT (*n*)	6	12	Sensitivity: 66.7

9.4.4 Language Support Group

As previously mentioned in Chapter 8 (Section 8.5.2.3), a group of children with language concerns (LS group) but who had not been evaluated by SLPs were recruited. The purpose was to determine whether the DA of word learning skills could accurately classify Singaporean bilingual preschool children with language concerns (LS group): those likely at risk of LI and require further evaluation versus those with language difference and were not at risk LI. This was achieved by comparing their performance on the DA of word learning skills with the recommendations in their end-of-intervention reports.

The LS group consisted of nine children with language concerns (six males, three females). Their mean age was 6;0 (years; months, *SD* = 6.09) and six participants were MD while the other three were ED (Table 9-19).

Table 9-19. Demographic information of the nine LS participants.

Participant No.	Age	Language dominance	Gender	Years attending local preschool	Housing (SES)	Mother's highest qualification
007	5;2	Mandarin	Female	1	HDB four-room	Secondary
039	5;9	English	Female	1	HDB four-room	Diploma/Professional qualification
086	5;5	English	Male	Not stated	HDB four-room	Primary
004	5;11	Mandarin	Male	3	HDB four-room	Diploma/Professional qualification
006	6;7	Mandarin	Male	4	HDB four-room	University
018	6;0	Mandarin	Male	4	HDB three-room	Diploma/Professional qualification
020	6;7	Mandarin	Male	4	HDB three-room	University

029	6;6	English	Male	3	Private	Diploma/Professional qualification
070	6;2	Mandarin	Female	4	HDB four-room	University

The recommendations by LSEds and progress made by all nine participants during the learning support sessions were retrieved from their end-of-intervention reports (See Table 9.20). Out of the nine participants, three participants made minimal progress and were recommended for another package of learning support intervention and/or further evaluation by SLP. A total of six LS participants were discharged as they had made considerable progress and were not observed to be at risk of any language or learning difficulties.

Based on their performance on the CELP P2 EV subtest, eight of the nine LS participants performed below average (scaled score of 6 and below) and were considered at risk of LI (i.e. SLT group) (see Table 9-20). The three participants (participant 007, participant 039, and participant 086) who were recommended for another package of intervention obtained below average scaled scores on the CELF P2 EV subtest. However five out of six of the participants (participant 004, participants 006, participant 018, participant 020, and participant 070) who made considerable progress and were recommended for discharged from DSP services also obtained below average scaled scores on the CELF P2 EV subtest. Therefore, based on their scores on the CELF P2 EV subtest alone, classification accuracy was at 44.4%

The logistic model was re-run to include the LS participants to obtain their predicted probabilities and predicted group membership based on their scores on the DA of word learning skills (Table 9-20). The six participants who made considerable progress and discharged were all predicted to be likely belong to the TD group based on their performance on the DA of word learning skills. This prediction corresponded to their progress in intervention and recommendations from LSEds (discharged from SLP services). However, two out of three

participants who were recommended for another package of intervention were predicted to belong to the SLT group with one participant (participant 007) being predicted to belong to the TD group. Classification accuracy of all nine LS participants based on their scores on the DA of word learning skills was 88.9%.

Table 9-20. Performance of all nine LS participants on the CELF P2 EV subtest, end-of-intervention recommendations and prediction of group membership based on their performance in the DA of word learning skills.

Participant No.	Recommendation from end-of-intervention report	CELF P2 EV subtest (based on scaled score)	Model prediction (based on scores obtained on the DA of word learning skills)
007	Referred for another package of intervention and evaluation by SLP	Below average	TD
039	Referred for another package of intervention and evaluation by SLP	Below average	SLT
086	Referred for another package of intervention and evaluation by SLP	Below average	SLT
004	Discharged	Below average	TD
006	Discharged	Below average	TD
018	Discharged	Below average	TD
020	Discharged	Below average	TD
029	Discharged	Low average	TD
070	Discharged	Below average	TD

9.5 Chapter Summary

The hypotheses presented in Chapter 8 (section 8.2) are now revisited to evaluate and summarise the results of the study.

Hypothesis 1:

1. The scores obtained on the DA of word learning skills by typically developing Singaporean bilingual children and Singaporean bilingual children who are identified as requiring on-going speech and language therapy can be differentiated. Specifically:

a) Singaporean bilingual children who are identified as requiring on-going speech and language therapy will require more mediation prompts to identify and name all four target words in the 'Teach' phase on the DA of word learning skills compared to typically developing Singaporean bilingual children.

b) Singaporean bilingual children who are identified as requiring on-going speech and language therapy will identify and name fewer target words in 'Retest' phase on the DA of word learning skills compared to typically developing Singaporean bilingual children.

Results:

- Results showed that SLT participants required more mediation prompts than TD participants to identify and name all four target words in the 'Teach' phase. The between-group differences were significant for TEACH Identification score and TEACH Naming score.
- Results showed that SLT participants identified and named fewer target words compared to TD participants in the 'Retest' phase. The between-group differences were significant for RETEST Identification score and RETEST Naming score.

Hypothesis 2:

2. The DA of word learning skills can evaluate the language skill of Singaporean bilingual children regardless of their prior language experience more accurately compared to a standardised assessment of vocabulary.

Specifically:

a) On the DA of word learning skills, children with the same developmental profiles regardless of their language experience will obtain similar patterns of

scores. Their performance on the DA of word learning skills is less likely to be influenced by their prior language experience.

b) On the standardised assessment of vocabulary, children's performance will be influenced by their prior language experience. Therefore, children with the same language developmental profiles but have different language experience will obtain a different range of scores.

Results:

- Results showed that regardless of their dominant language, TD MD participants and TD ED participants obtained similar scoring patterns on the DA of word learning skills, no statistical group differences in scores were observed. Likewise, SLT ED and SLT MD participants obtained similar scoring patterns on the DA of word learning skills, no statistical group differences in scores were also observed.
- However, on the CELF P2 EV subtest, between-group differences (TD MD and TD ED; SLT MD and SLT ED) were observed.

Hypothesis 3:

3. The set of 'Teach' (TEACH Identification, TEACH Naming) and 'Retest' (RETEST Identification, RETEST Naming) scores obtained on the DA of word learning skills can classify Singaporean bilingual children who are referred and identified to require on-going speech and language therapy and those who were typically developing with high accuracy. The classification accuracy of the DA of word learning skills will be at least 80.0% (Sensitivity: 80.0%, Specificity: 80.0%), demonstrating satisfactory classification accuracy (DiStefano & Morgan; 2011; Glascoe, 2005).

Results:

- Results showed that the DA of word learning skills accurately classified 93.9% of the participants (Sensitivity: 88.9%; Specificity 95.8%). The CELF P2 EV subtest, however, only accurately classified 74.2% of the participants (Sensitivity: 66.7%; Specificity: 77.1%).

Hypothesis 4:

4. The performance of Singaporean preschool bilingual children with language concerns on the DA of word learning skills can differentiate those who are 1) likely at risk of LI and will require further evaluation, and 2) likely demonstrating language difference (i.e. typically developing) and will not require further evaluation.

   Results:

   - A group of nine participants with language concerns (LS group) who received language learning support services in the DSP were evaluated on the DA of word learning skills. Their performance on the DA of word learning skills was compared to the recommendations (i.e. continue another package of language support services and further evaluation; or discharged from DSP, not at risk of LI) provided by LSEds in their end-of-intervention reports. Out of the nine LS participants, eight of them (88.9%) obtained scores on the DA of word learning skills that corresponded to the recommendations in their end-of-intervention reports.

In the next chapter, the results are further elaborated upon, with reference to the evidence to be gathered under the Validity Argument Framework (Kane, 2006; 2013).

Chapter 10 Phase Two: Discussion

10.1 Introduction

The primary objective of Phase Two of the study was to evaluate whether a DA process, specifically a DA of word learning skills, could validly evaluate the language skills of bilingual preschool children, and therefore identify those at risk of LI. In this current study context, the performances of English-Mandarin speaking bilingual preschool children who were referred for on-going speech and language therapy (SLT group) and their typically developing counterparts (TD group) on the DA of word learning skills were evaluated and compared. Specifically, this study investigated whether 'Retest' scores (i.e. RETEST Identification, RETEST Naming scores), combined with the children's modifiability scores (i.e. TEACH Identification, TEACH naming scores) could differentiate children who were typically developing (i.e. TD group) and children who likely had LI (i.e. SLT group). It was hypothesised that the scores obtained on the DA of word learning skills would 1) differentiate the language learning potential between both groups, and 2) classify the children into respective groups (i.e. SLT group, TD group) with higher accuracy compared to a standardised assessment of vocabulary.

The children's word naming, word identification and modifiability scores were evaluated in a 'Test-Teach-Retest' DA paradigm. The target words were four non-words that were associated with unfamiliar object referents. Unlike previous studies (e.g. Kapantzoglou et al., 2012; Petersen et al., 2017; Peña et al., 2001; 2006; Ukrainetz, Harpell, Walsh & Coyle, 2000) where examiners' ratings of the children's modifiability were evaluated subjectively on Likert-scale checklists, this study used a GP framework to objectively examine the children's modifiability by evaluating the level of mediation prompts provided in 'Teach' phase to identify and name all four target words (TEACH identification score, TEACH naming score). For word identification, it was hypothesised that the SLT participants would require more mediation

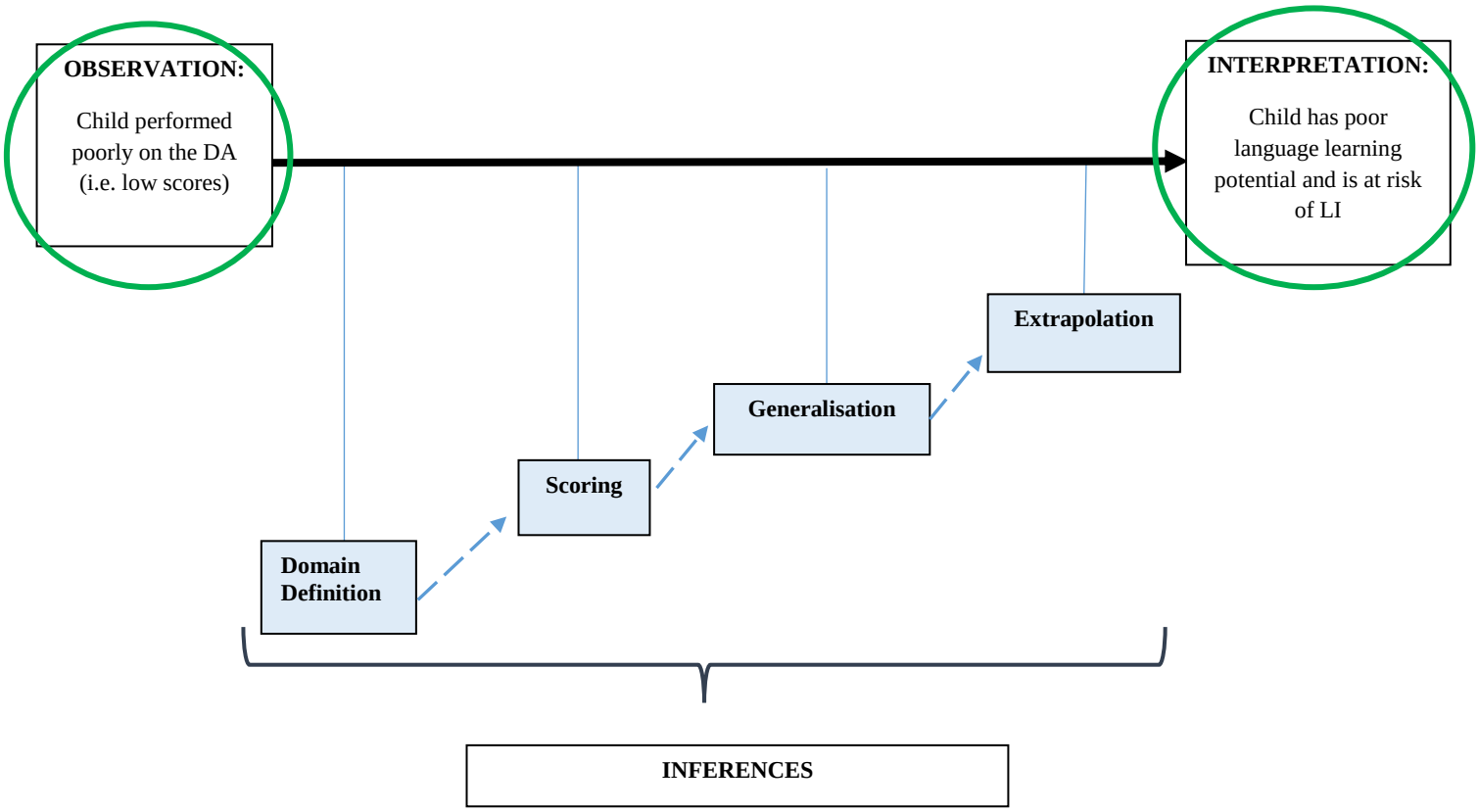
prompts in the 'Teach' phase and would identify fewer target words in the 'Retest' phase. For word production, it was also hypothesised that the SLT participants would require more mediation prompts in the 'Teach' phase and would name fewer target words in the 'Retest' phase.

10.2 The Validity Argument Framework

To explore whether the DA of word learning skills could accurately and validly differentiate English-Mandarin speaking bilingual preschool children who were referred for ongoing speech and language therapy (SLT group) from their typically developing counterparts (TD group) and therefore identify bilingual preschool children at risk of LI, the Validity Argument Framework (Kane, 2006; 2013) was adopted in this study to develop and evaluate the evidence. The validation process began by stating the proposed interpretations and use of the DA. When a child is observed to score poorly on the DA, the child is likely to have poor language learning ability and is at risk of LI (Figure 10-1, circled in green).

The development of the DA of word learning skills and the result of the children's performance on the DA is discussed as part of the evidence to support claims made under each inference of the Validity Argument Framework (Kane, 2006; 2013; Figure 10-1). In addition, when appropriate, the children's performance on a standardised assessment of vocabulary (i.e. CELF P2 EV subtest) is compared to further substantiate the evidence for the DA of word learning skills. As mentioned previously (Chapter 4, Section 4.7), during the development and exploration of a new assessment tool/approach (i.e. DA of word learning skills), the evidence gathered is of confirmationist bias. This is to advocate and support the proposed use and interpretation of the DA as an appropriate alternative assessment approach to evaluate the language learning potential of bilingual children and therefore identify bilingual children at risk of LI.

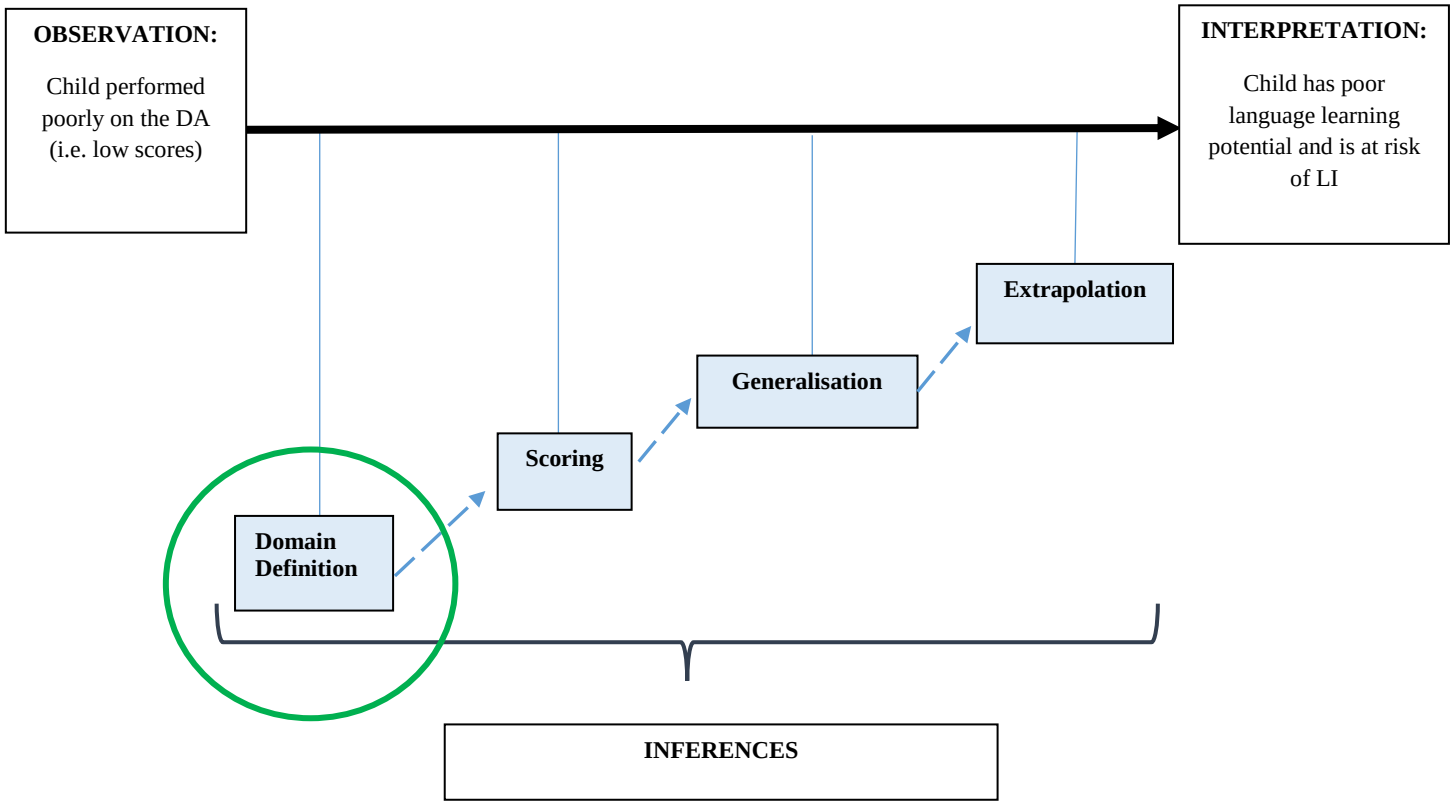
Figure 10-1. Interpretation and use of the DA.



10.2.1 Domain Definition: Evaluating the Evidence

The first chain of evidence to be gathered was to support the claim under *Domain Definition* of the Validity Argument Framework (Kane, 2006; 2013): that the observations of performance on the DA reveal a preschool child's language learning potential, and this can be used to determine whether the child is at risk of LI (Figure 10-2, circled in green). The assumption was that the DA task chosen is developmentally appropriate and least biased for evaluating language learning potential among bilingual children. Support for this claim came from the review of the literature to identify a suitable language task during the development of the DA process (Chapter 5).

Figure 10-2. Evaluating the evidence for Domain Definition.



To recap, the theoretical accounts of LI (i.e. linguistically-based accounts and processing-based accounts) were reviewed to identify a suitable language task to be assessed in the DA context. The language task selected choosen had to be one where poor performance was a behavioural indicator of LI. Under the linguistically-based accounts, it is believed that the main language deficits are in the children's underlying grammar (Bedore & Peña, 2008). However, the acquisition and mastery of verb tense marking among bilingual children occurs at a different rate and developmental trajectory compared to monolingual developmental norms. This difference in acquisition and mastery of verb tense marking is secondary to the variability in language experience among bilingual children. Thus, grammatical-type tasks under the linguistically-based accounts are not appropriate to be evaluated in the DA context for the identification of LI among bilingual children.

The limitations in using linguistically-based behavioural indicators to identify LI among bilingual children have turned the focus towards processing-based behavioural indicators

instead. Under the processing-based account, it is assumed that language difficulties observed are secondary to limitations or deficits in cognitive processing capacity used for processing linguistic information as well as other cognitive operations (Bedore & Peña, 2008; Paradis, 2010). An advantage of using language processing tasks is that the influence of prior language experience can be minimised, especially in tasks where the language materials are novel. Moreover, evaluating children's underlying cognitive ability to process new language materials is parallel to the purpose of DA which is to evaluate children's language learning potential. This provides a more accurate assessment of their true language learning potential and therefore language ability. To evaluate the language processing ability of young bilingual children, word learning, a language processing task was identified as a developmentally appropriate task to be evaluated in a DA context.

Further evidence to support the choice of a word learning task, a processing based task, was evident from the significant difference in non-verbal intelligence scores obtained by the SLT and TD group on the PTONI. The PTONI was used as an identification measure in Phase Two of the study (Chapter 8, Section 8.6.3). The purpose was to ensure that all children who were included in the final data analysis presented with normal range non-verbal intelligence and that those who presented with below average intelligence (Non-Verbal Index Score <90) were excluded. This was in line with the general definition of LI described in the literature where language difficulty occurs in the absence of other developmental deficits (i.e. cognitive levels within normal limits) (Bishop, 2014; Leonard, 1998; Reilly et al., 2014).

Both the SLT group and TD group performed within the average to above average range on the PTONI (Non-Verbal Index Score > 90). However, the SLT group performed less well compared to the TD group. On average, the SLT group (M= 114.16) scored 12 standard points below the TD group (M= 126.03). This result coincides with the findings from other studies that low average non-verbal intelligence is part of the profile of children with LI (Gallinat &

Spaulding, 2014; Eare, Gallinat, Grela, Lehto & Spaulding, 2017). For example, Gallinat and Spaulding (2014) conducted a meta-analysis in which they compared the difference in non-verbal performance of children with and without LI on nonverbal cognitive tests across 131 studies. After adjusting for the differences in the tests used, they found that on average, children with LI scored slightly more than 10 standard points below their similarly-aged counterparts but still within the lower average range.

Fair to low average performance of children with LI on non-verbal intelligence tests is consistent with the processing-based account of LI. Children with LI have fair to low average non-verbal intelligence can present with slower ability in processing both linguistic and non-linguistic information compared to typically developing children. (Leonard et al., 2007). Therefore on a novel word learning task, children with LI are expected to have difficulties processing, storing and retrieving linguistic and non-linguistic information about the new words compared to typically developing children. However, this limitation in cognitive processing may only be obvious under low exposure conditions to the novel words (Chapter 5, Section 5.4.2). The limitations in cognitive processing observed in children with LI can be compensated under conditions of increased exposure to the novel words (Gray, 2003, Kapantzoglou et al., 2012). Therefore, to differentiate the performance the performance of typically developing bilingual children with and without LI on the DA of word learning skills, the novel words have to be exposed under the conditions of low exposure.

The fair to low average non-verbal intelligence performance of the SLT group further supports the selection of the word learning under low exposure (i.e. nine times) under the processing-based account of LI to identify bilingual children at risk of LI (Table 10-1, bolded).

Table 10-1. Evidence to support the claim under Domain Definition.

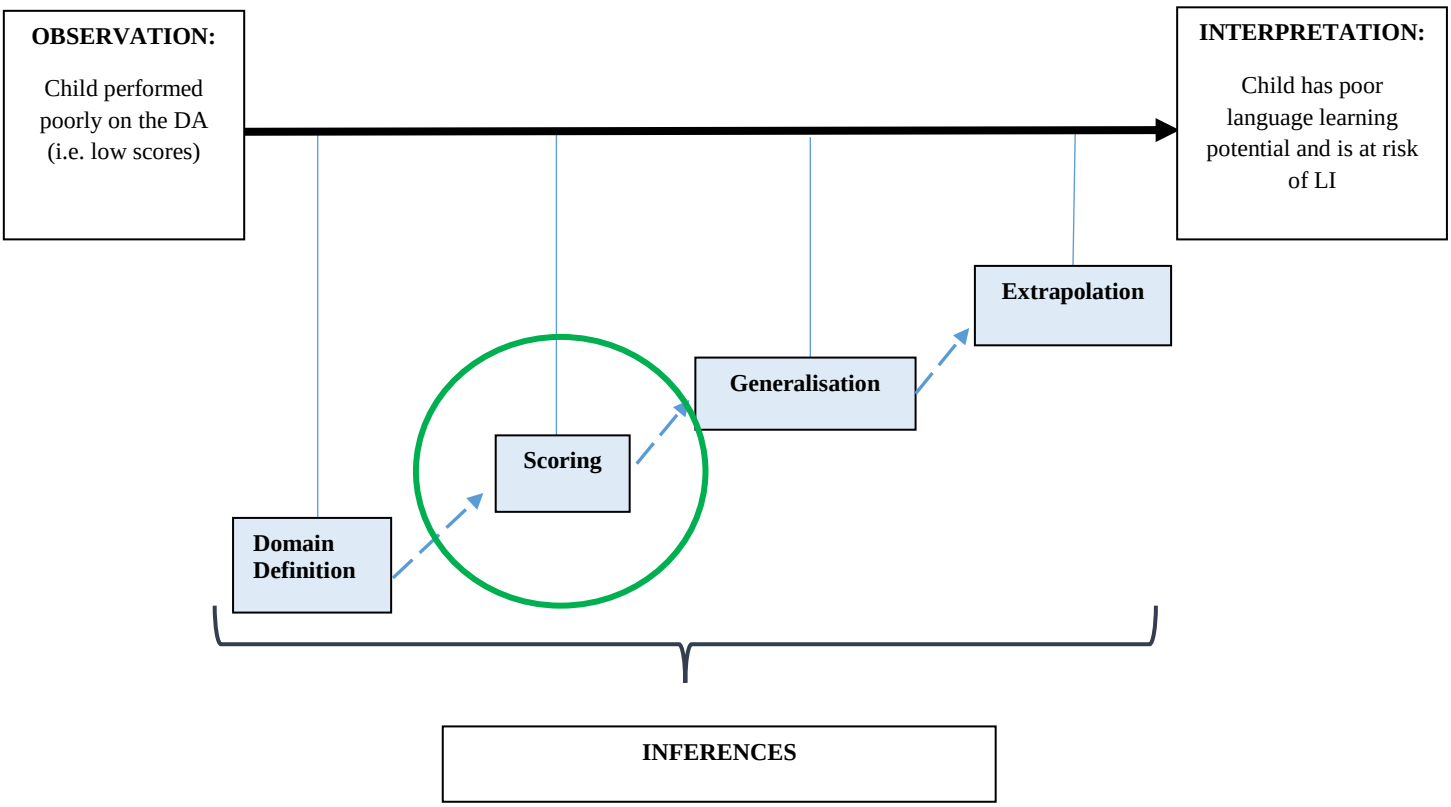
Inference	Claims	Assumptions	Evidence
Domain Definition	Observations of performance on the DA can reveal bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI.	The language task selected to be evaluated in the DA is developmentally appropriate and least biased for evaluating language learning potential among bilingual children.	A review of the literature identified language processing tasks under the processing-based account of LI to evaluate the language learning potential of bilingual children with least biased. This is because language processing tasks evaluate children's ability to learn new language materials where their performance is least influenced by their prior language experience. Word learning is identified as a developmentally appropriate processing-based task to evaluate the language learning potential of young bilingual children as vocabulary acquisition is one of the first sublinguistic domains that young children acquire. A review of the literature provided emerging evidence that bilingual children with LI performed poorly on word learning tasks compared to their typically developing counterparts under low exposure conditions. Therefore, poor performance on word learning tasks under low exposure conditions can be a behavioural indicator for LI and identify bilingual children at risk of LI.

			Results showed that SLT participants (i.e. children with LI) obtained lower non-verbal intelligence scores compared to TD participants (i.e. typically developing children). Fair to low average performance of children with LI on non-verbal intelligence tests is consistent with the processing-based account of LI. The difficulties in processing linguistic and non-linguistic information (i.e. word learning) can be attributed to their fair to low average cognitive processing abilities.

10.2.2 Scoring: Evaluating the Evidence

The next chain of evidence to be gathered was to move from a claim about the selection of a relevant task (i.e. word learning task) to be evaluated in a DA context to the application and accuracy of the method used to evaluate children's language learning potential (i.e. word learning ability) in a DA context (Chapelle & Voss, 2014). In other words, whether the task administrative conditions and scoring rubric of the DA were appropriately designed to evaluate the language learning potential of bilingual children (Figure 10-3, circled in green).

Figure 10-3. Evaluating the evidence for Scoring.



The evidence to support the claim (*Observations of performance on the DA task are evaluated to provide observed scores reflective of language learning potential*) under the *Scoring* was gathered in Chapter 6 during the development and piloting of the DA process. The assumptions were: 1) the task procedure provides opportunities for language learning; 2) the task administrative conditions are appropriate for observing and scoring language learning opportunities and; 3) the rubric for scoring responses are appropriate for providing evidence of language learning ability.

The summary of evidence to support the claims under *Scoring* is represented in Table 10-2. To recap and summarise, the design of the DA of word learning skills was put together after a review of the literature. Firstly, there was a careful selection of the stimuli. Four non-words with high phonotactic probability and low neighbourhood density were selected as target words as these word characteristics were found to maximise the difference in word learning ability

between children with and without LI (Chapter 6; Section 6.2.3). Next, the procedure for administering the teaching instructions in a 'Test-Teach-Retest' format was designed to be structured and scripted to ensure consistency of administrative procedures across all children (Chapter 6, Section 6.2.5). In addition, each target word was presented only nine times (i.e. low exposure) in the interactive play activity in Stage 2 of the 'Teach' Phase (Chapter 6; Section 6.2.5.3). This ensured that sufficient opportunities were provided for word learning to take place among typically developing children while the word learning task remained challenging for children with LI under the condition of low exposure. Last but not least, to ensure that there was objective scoring of the children's performance on the DA of word learning skills: 1) A GP framework was used to score children's modifiability (i.e. based on the level of mediation prompts provided) in response to the teaching instructions provided in the 'Teach' phase, and 2) Children's ability to learn all four target words were evaluated in the 'Retest' phase based on the number of target words they could identify and recall independently.

A pilot study was then conducted to trial the DA of word learning skills (Chapter 6, Section, 6.3). After minor revisions were made, a second pilot study confirmed that 1) the DA procedure provided appropriate opportunities within the assessment itself for word learning to take place as both typically developing children and children with LI were able to learn the target words within their ability, and 2) the scoring rules could be applied objectively to accurately differentiate the scores (i.e. based on the level of mediation prompts provided in the 'Teach' phase; number of target words identified or named in the 'Retest' phase) obtained by typically developing bilingual children and bilingual children with reported language difficulties, reflecting their differences in both modifiability and language learning ability.

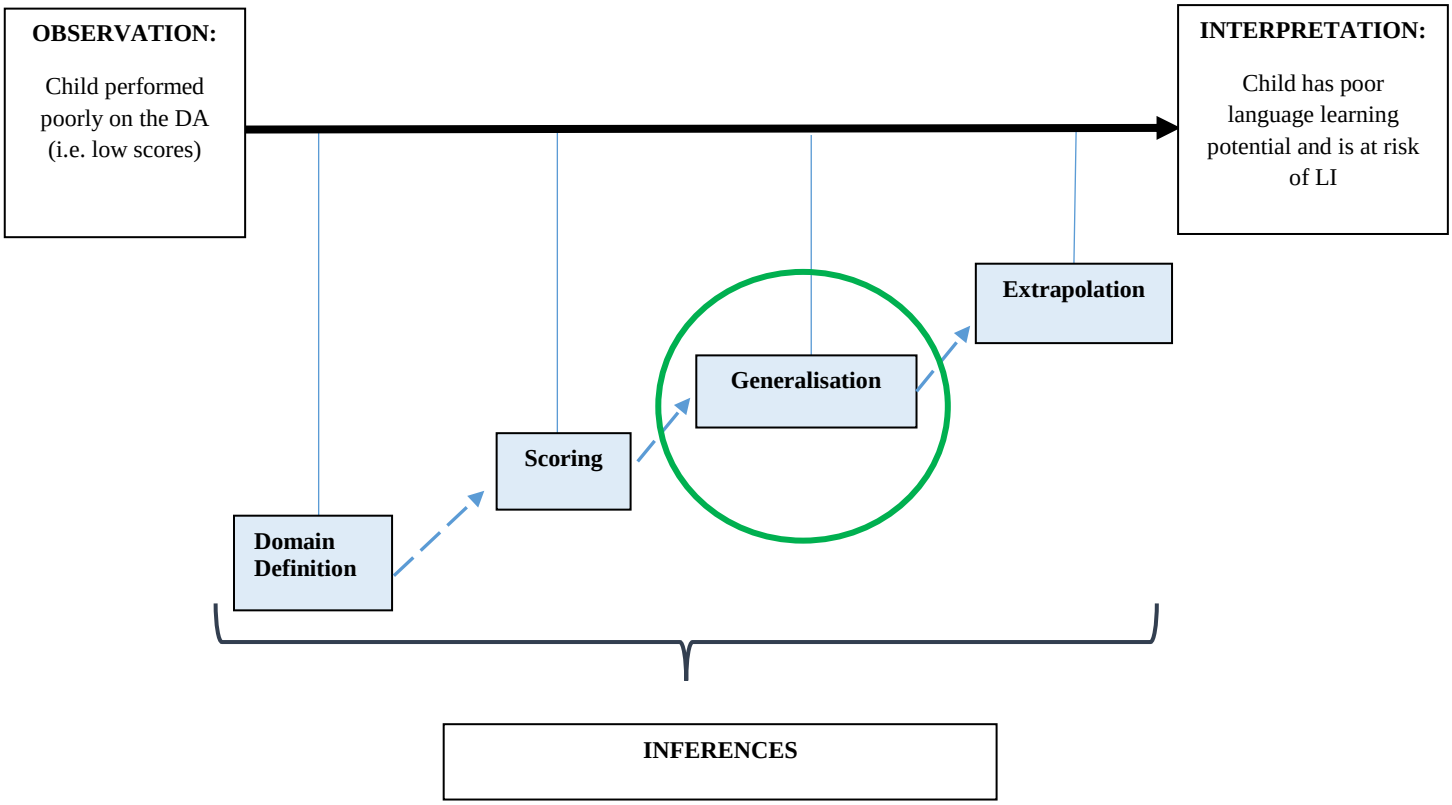
Table 10-2. Evidence to support the claim under Scoring.

Inference	Claims	Assumptions	Evidence
Scoring	Observations of performance on the DA task are evaluated to provide observed scores reflective of language learning potential.	The task procedure provides opportunities for language learning.	Opportunities were created in the 'Teach' Phase for the children to learn all four target words.
		The task administrative conditions are appropriate for observing and scoring language learning opportunities.	The administrative procedure of the DA was designed to be structured and scripted. This ensured that the learning conditions across children were consistent and optimum for word learning to take place.
			There was systematic development of the rubric for scoring children's modifiability (i.e. hierarchy of mediated prompts provided in the 'Teach' phase) and language learning potential (i.e. number of target words identified and named correctly in the 'Retest' phase).
		The rubric for scoring responses are appropriate for providing evidence of language learning ability.	Results from the pilot study showed that both typically developing bilingual children and children with reported language difficulties were able to demonstrate evidence of language learning. There was a difference in all four scores obtained between typically developing children and children with LI.

10.2.3 Generalization: Evaluating the Evidence

The next chain of evidence to be gathered was to demonstrate that the DA task is able to capture accurate, consistent and relevant summaries of children's language learning ability to make the next claim under *Generalisation* (Chapelle & Voss, 2014). Specifically, the observed scores on the DA reflect language learning potential that are consistent among bilingual children with the same language developmental profiles regardless of their prior language experience (Figure 10-4, circled in green). The assumptions were that: 1) the scores obtained on the DA by children with and without LI (i.e. SLT group and TD group) are differentiated, and 2) The scores obtained on the DA by bilingual children with the same language developmental profiles are consistent regardless of their prior language experience (i.e. TD ED participants and TD MD participants).

Figure 10-4. Evaluating the evidence for Generalisation.



To gather the evidence, the following results were analysed in the following order:

1. The scores obtained on the DA of word learning skills by SLT and TD participants.
2. The scores obtained on the DA of word learning skills by participants with the same language developmental profiles but different language experience (i.e. TD ED participants versus TD MD participants; SLT ED participants versus SLT MD participants).

The performances of SLT participants and TD participants on the DA of word learning skill are compared and elaborated below (Section 10.2.3.1). In addition, to demonstrate that the DA of word learning skills is a less biased assessment approach, the participants' performance on the CELF EV subtest, a standardised vocabulary, will also be compared (Section 10.2.3.2).

10.2.3.1 Dynamic Assessment of Word Learning Skills

The results found significant group differences between SLT and TD group in all four scoring aspects of the DA of word learning skills. For word identification, the SLT group required more mediation prompts to identify all four target words in the 'Teach' phase (TEACH Identification score: SLT group Mean = 4.741 versus TD group Mean = 4.104) and identified fewer target words in the 'Retest' phase (RETEST Identification score: SLT group Mean = 1.875 versus TD group Mean = 3.872). For word naming, the SLT group required more mediation prompts to identify all four target words in the 'Teach' phase (TEACH Naming score: SLT group Mean = 14.319 versus TD group Mean = 12.601) and named fewer target words in the 'Retest' phase (RETEST Naming score: SLT group Mean = 0.355 versus TD group Mean = 2.102).

The results obtained can be interpreted and supported theoretically in terms of word learning ability. Based on the processing based account of LI, children with LI have limitations in their processing capacity that hinder their ability establish semantic and phonological information between a novel word and the object referent under conditions of low exposure (Alt et al., 2004,

Ellis Weismer & Evans, 2002, Kan & Windsor, 2010). This result is consistent with the findings from other studies (Gray, 2003; Gray, 2005; Rice, Oetting, Marquis, Bode & Pae, 1994) where children with LI required twice as much input (i.e. more exposure, more encoding cues) from the examiner to facilitate word learning. Therefore as expected, the SLT participants compared to TD participants required more mediation prompts to establish novel word – object referent during identification while requiring the almost maximum number of mediation prompts to name all four target words in the 'Teach phase'.

This result was consistent with the findings from Kapantzoglou and colleagues study's (2012) where they found that typically developing children made more associations between the phonological and semantic representation of new words compared to children with LI in a brief DA learning context. However, in Kapantzoglou's study (2012), group difference was only found for word identification and not for word naming in the 'Retest' phase. Due to methodological differences of the DA used in the current study and Kapantzoglou's study, it was likely that the additional mediation prompts provided in Stage 1 and Stage 3 of the 'Teach' phase in the current study emphasised the phonological and semantic representations of the target words. Therefore, the additional mediation prompts likely assisted TD participants in establishing stronger representations of the target words. Consequently, this resulted in TD participants naming more target words in 'Retest' phase. Whilst due to the limitations in processing capacity among SLT participants, the mediation prompts provided likely had a minimal to no effect on the word learning performance for SLT participants.

The performance on the DA of word learning skills within each group (TD group, SLT group) were further analysed to determine if prior language experience (i.e. language dominance) had any impact on their performance on the DA of word learning skills. Within the TD group, there was no significant group difference (i.e. ED versus MD) in the number of mediation prompts required in the 'Teach' phase to identify and name target words. Similarly, there was

no significant difference in the number of target objects that were successfully identified and named in the 'Retest' phase between TD ED and TD MD children. The same results were also observed in the SLT group where there was no group difference for both SLT ED children and SLT MD children in the number of mediation prompts required in the 'Teach' phase and the number of target words identified and named in the 'Retest' phase. The findings suggest that regardless of the children's dominant language, the scores obtained on the DA of word learning skills were able to capture the word learning profiles of the children based on their language developmental profiles (i.e. SLT vs TD). In other words, the participants' prior language experience did not impact on their performance on the DA of word learning skills.

10.2.3.2 Clinical Evaluation of Language Fundamentals Preschool 2 Expressive Vocabulary Subtest

The CELF P2 EV subtest is a common language assessment that is used to assess the end-product of word learning (i.e. lexical knowledge). Unlike the DA of word learning skills, the CELF P2 EV subtest is a standardised assessment that evaluates children's English lexical knowledge at the point of assessment without evaluating their ability to learn. It does not take into consideration whether the child has the relevant language experience or language exposure to the English language. In other words, it does not provide a fair evaluation of the lexical ability of bilingual children as performance is dependent on the amount of previous exposure each child has to the English language. Furthermore, the stimuli and pictures used in the CELF P2 EV subtest were designed to be familiar to monolingual English-speaking children from the USA. Therefore, the stimuli and pictures in the CELF P2 EV subtest may not be culturally and linguistic appropriate or familiar to bilingual children who have a different cultural and linguistic experience. However, in spite of its limitations, many SLPs are still using standardised assessment of vocabulary such the CELF P2 EV subtest to evaluate bilingual children's vocabulary knowledge (Teoh et al., 2012; 2017).

Although the SLT group performed significantly more poorly than the TD group on the CELF P2 EV subtest, the covariate – language dominance was found to have a significant influence on the children's performance on the CELF P2 EV subtest. Further analysis showed that within the TD group, TD MD children performed significantly more poorly than their TD ED counterparts even when both groups of children were identified to have the same language developmental profiles (i.e. typically developing). All TD ED participants obtained a scaled score of 7 and above which indicated average expressive vocabulary ability. However, for TD MD participants, 11 out of 29 obtained a scaled score of 6 and below, indicating below average expressive vocabulary abilities. In another study, Teoh and colleagues (2012) reported a similar result where 21 out of 41 typically developing Mandarin dominant bilingual children obtained a scaled score of 6 and below on the CELF P2 EV subtest. The findings from Teoh et al. (2012) and the current study clearly demonstrate that using a standardised assessment of vocabulary to assess the English vocabulary skills of non-English dominant bilingual children may result in assessment bias. This may lead to an over-identification of typically developing non-English dominant bilingual children at risk of LI when they, in fact, have normal language ability.

Similarly, for the SLT group, the SLT MD participants performed significantly more poorly than the SLT ED counterparts on the CELF P2 EV subtest. Camilleri and Law (2007) also reported in their study that children with English as an additional language who were receiving speech and language therapy obtained significantly lower scores on the British Picture Vocabulary Scale II (Dunn et al., 1997), a standardised assessment, compared to monolingual English-speaking children who were receiving speech and language therapy. The findings from the current study and Camilleri's study suggest that assessing non-English speaking bilingual children with LI on English standardised assessments may, in fact, overestimate the severity of their language difficulties.

10.2.3.3 Which is a Less Biased Assessment Measure?

The performance of bilingual children on the DA of word learning skills was reflective of their language developmental profile (i.e. SLT participants, TD participants) regardless of their prior language experience. SLT participants, irrespective of their dominant language, performed less well on all scoring components of the DA of word learning skills compared to TD participants. The pattern of scores obtained by both SLT ED participants and SLT MD participants on all four scoring components of the DA of word learning skills was comparable as there was no significant group difference found between both groups. Likewise, the pattern of scores obtained by TD ED and TD MD children on all four scoring components on the DA of word learning skills were comparable with no significant group difference found between both groups. However, on the CELF P2 EV subtest, a frequently used standardised assessment used by SLPs in Singapore to assess the expressive vocabulary skills of Singaporean children, the participants' performance was influenced by their prior language experience.

The comparison of each group's performance on the DA of word learning skills and CELF P2 EV subtest clearly identified DA of word learning skills as a less biased measure of bilingual children's lexical ability. This finding provides preliminary evidence under *Generalisation* of the Validity Argument Framework (Kane, 2006; 2013) that the scores obtained by children on the DA of word learning skills were consistent among bilingual children with the same language developmental profiles (Table 10-3). Specifically: 1) bilingual children identified for on-going speech and language therapy (SLT group) on average required more mediation prompts in the 'Teach' phase to identify and name all four target words compared to their typically developing counterparts (TD group) regardless of their prior language experience, and 2) bilingual children identified for on-going speech and language therapy (SLT group) on average identified and named fewer target words in the 'Retest' phase compared to their

typically developing counterparts (TD group) regardless of their prior language experience (Table 10-3).

Table 10-3. Evidence to support the claim under Generalisation.

Inference	Claim	Assumptions	Evidence
Generalisation	Observed scores on the DA are consistent among bilingual children with the same language developmental profiles regardless of their prior language experience.	The scores obtained on the DA between bilingual children with and without LI are differentiated.	Bilingual children identified for on-going speech and language therapy (SLT group) on average required more mediation prompts in the 'Teach' phase to identify and name all target words compared to their typically developing counterparts (TD group). Bilingual children identified for on-going speech and language therapy (SLT group) on average identified and named fewer target words in the 'Retest' phase compared to their typically developing counterparts (TD group).
		The scores obtained on the DA by bilingual children with the same language developmental profiles are consistent regardless of their prior language experience.	There were no group differences observed in the scores obtained between typically developing bilingual children with variable language experience (TD ED participants versus TD MD participants). There were no group differences observed in the scores obtained between bilingual children identified for on-going speech and language therapy with variable language experience (SLT ED participants and SLT MD participants).

10.2.4 Extrapolation: Evaluating the Evidence

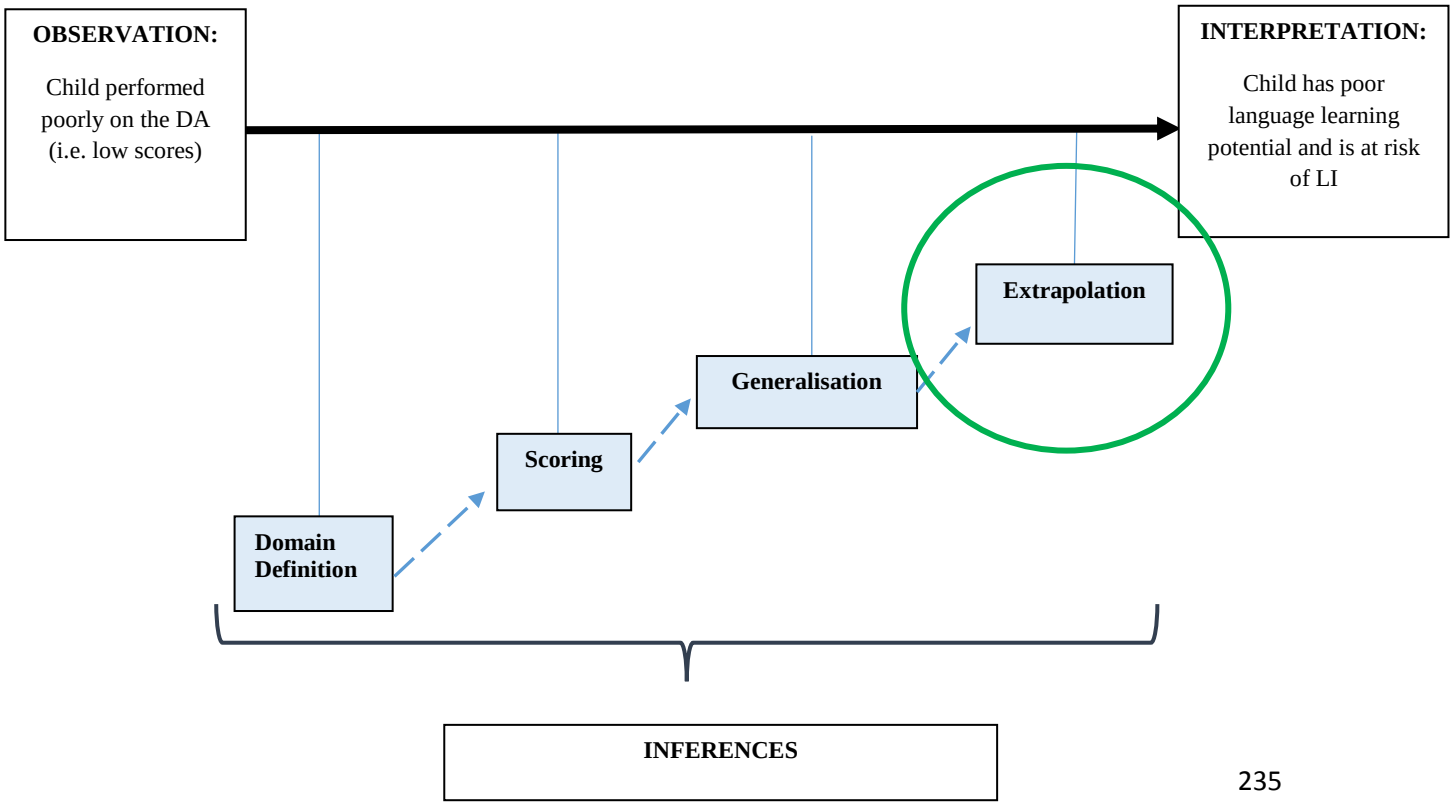
The final chain of evidence to be gathered was to consider the correlation between the scores obtained by both SLT and TD participants on the DA of word learning skills and their language

developmental profiles in the 'real' world (Chapelle & Voss, 2014; Cook et al., 2015). This was to support the claim under *Extrapolation*; that the construct of language learning potential as assessed on the DA predicts scores that correspond to the bilingual children's language developmental profiles (Figure 10-5). The assumption was that the performance on the DA could predict the bilingual children's language developmental profiles (i.e. based on group membership: SLT group versus TD group).

To gather the evidence, the following results were analysed in the following order:

1. The classification accuracy of the DA of word of learning skills.
2. The classification accuracy of the DA of word learning skills versus the classification accuracy of CELF P2 EV subtest.
3. The performance of bilingual children with language concerns (LS group) on the DA of word learning skills, and whether their scores corresponded to the recommendations in their end-of-intervention reports.

Figure 10-5. Evaluating the evidence for Extrapolation.



10.2.4.1 Classification Accuracy of the Dynamic Assessment of Word Learning Skills

Binomial logistic regression was performed to evaluate how accurately the participants' performance on the DA of word learning skills predicted their group memberships (SLT group or TD group). The analysis was conducted using all four scoring components (TEACH Identification, TEACH Naming, RETEST Identification, RETEST Naming) where between-group differences were found (as discussed in section 10.2.3.1). Overall, the model classified 93.9% of the participants accurately, with 88.9% sensitivity and 95.8% specificity. The results suggest that the DA of word learning skills demonstrated satisfactory classification accuracy (DiStefano & Morgan; 2011; Glascoe, 2005) as an alternative assessment approach.

The high classification accuracy achieved by the DA of word learning skills developed and explored in the current study coincided with results achieved by other DA studies in the field of language testing (i.e. Kapantzoglou et al., 2012; Peña et al., 2001, 2006, 2017; Petersen et al., 2017). In other studies, it was found that modifiability scores/ratings combined with 'Retest' scores consistently achieved high classification accuracies (i.e. Kapantzoglou et al., 2012; Peña et al., 2001, 2006, 2017; Petersen et al., 2017). However, unlike previous studies where children's modifiability was rated subjectively on Likert Scale Checklists (i.e. MS and LSC), the DA of word learning skills used a GP framework where children's modifiability was objectively scored by recording the total number of mediation prompts provided by the examiner in the 'Teach' phase.

To the researcher's knowledge, there were two other DA studies (Camilleri & Law, 2007; Hasson et al., 2012) that also adopted a GP framework to score children's modifiability in learning new words. Although findings from both studies found that that the number of graduated prompts needed to learn new words was differentiated between children with and without LI, with children with LI requiring more prompts, the classification accuracies of the word learning tasks in both studies were not evaluated or reported. Therefore, the findings from

the current study are the first in the field to demonstrate that DA using a GP framework to score children's modifiability in the 'Teach' phase combined with 'Retest' scores can result in high classification accuracy. The advantages of an objective quantifying measurement to evaluate modifiability will be elaborated on later in the discussion (Section 10.3).

10.2.4.2 Classification Accuracy of the Clinical Evaluation of Language Fundamental Preschool 2 Expressive Vocabulary Subtest

Overall, only 74.2% of the participants were classified correctly based on their scores obtained on the CELF P2 EV subtest. On the CELF P2 EV subtest, a scaled score of 6 and below indicates that the child performed below average when compared to same-age peers and is at risk of LI. Specificity was calculated to be 77.1 % as there were 11 TD participants (out of 48) who performed below average and were identified at risk of LI. On further analysis, all the 11 TD participants were MD, providing further evidence that standardised assessments such as the CELF P2 EV subtest constitute biased assessment for evaluating the lexical skills of bilingual children from non-English dominant backgrounds. Sensitivity was calculated to be at 66.7% where only 12 out of 16 SLT participants performed below average and were identified at risk of LI. Fair to poor sensitivity and specificity percentages obtained suggest that the CELF P2 EV subtest did not demonstrate satisfactory classification accuracy (DiStefano & Morgan; 2011; Glascoe, 2005) to identify bilingual children at risk of LI.

10.2.4.3 Classification Accuracy: Dynamic Assessment of Word Learning Skills versus Clinical Evaluation of Language Fundamentals Expressive Vocabulary Subtest

The findings showed that the DA of word learning skills improved the classification accuracy of bilingual children especially if the child is from a non-English dominant background. In other words, compared to a standardised assessment of vocabulary, the occurrence of test bias (i.e. cultural and linguistic bias) was minimised. This finding provides preliminary evidence

for *Extrapolation* that performance on the DA of word learning skills can discriminate between bilingual preschool children who were referred for ongoing speech and language therapy (SLT group) from their typically developing bilingual counterparts with high classification accuracy (>80.0%) compared to a standardised assessment of vocabulary (<80.0%).

10.2.4.4 Classification Accuracy of the Dynamic Assessment of Word Learning Skills for Learning Support Group

Nine English-Mandarin bilingual children (LS group) who were identified with language concerns but not evaluated by SLPs were also assessed on the DA of word learning skills and the CELF P2 EV subtest. The purpose was to determine whether the DA of word learning skills could accurately differentiate and classify this group of participants with language concerns into those likely at risk of LI and requiring further evaluation, and those with language difference that do not require further evaluation. This was achieved by comparing their performance on the DA of word learning skills with their end-of-intervention recommendations.

Out of the nine LS participants, reviews of their end-of-intervention records by the LSEds showed that six of them made good progress during the language support sessions conducted. The LSEds in consultation with a SLP had recommended for all six of them to be discharged from DSP services. In other words, these six LS participants were observed to present with language difference and not at risk of LI. Similarly, on the DA of word learning skills, these six LS participants were predicted by the logistic model to belong to the TD group as they obtained scores that were comparable to the TD participants. However, on the CELF P2 EV subtest, five out of six of them obtained a scaled score of 6 and below which indicated that they had below average expressive vocabulary skills and were at risk of LI.

Three out of the nine LS participants made minimum progress during the language learning support sessions conducted. They were recommended in their respective end-of-intervention

reports to continue a second package of language learning support and further evaluation. In other words, these three participants were observed to be likely at risk of LI. These three participants also obtained a scaled score of 6 and below on the CELF P2 EV subtest, indicating below average expressive language skills. However, on the DA of word learning skills, only two out of the three participants were predicted by the logistic model to belong to the SLT group while one participant was predicted to belong to the TD group. Therefore, there was one child (Participant 007, Table 10-4) whose performance on the DA of word learning skills did not correspond to the recommendations made by the LSEd in the end-of-intervention report.

To summarise, based on scores obtained by the LS participants on the DA of word learning skills, the model prediction for eight out of the nine children corresponded to the recommendations obtained in their end-of-intervention reports (Table 10-4). While based on their performance on the CELF P2 EV subtest, only four out of nine LS participants obtained scores on the CELF P2 EV subtest that corresponded to the recommendations obtained in their end-of-intervention reports (Table 10-4, those that corresponded are highlighted in green, those that did not match are highlighted in red).

Table 10-4. Matching the LS participants' performance on the DA of word learning skills and CELF P2 EV subtest to the recommendations from their end-of-intervention reports.

Participant No.	**Recommendations from end-of-intervention reports**	**CELF P2 EV subtest (based on scaled score)**	**Model prediction (based on DA of word learning skills scores)**
007	Referred for another package of intervention and evaluation by SLP	Below average	TD
039	Referred for another package of intervention and evaluation by SLP	Below average	SLT
086	Referred for another package of intervention and evaluation by SLP	Below average	SLT
004	Discharged	Below average	TD
006	Discharged	Below average	TD
018	Discharged	Below average	TD
020	Discharged	Below average	TD

029	Discharged	Low average	TD
070	Discharged	Below average	TD

Remarks: Those that corresponded are highlighted in green, those that did not are highlighted in red

There were, however, methodological limitations with the group of LS participants that may hinder the interpretations of the above findings. Firstly, LS participants were not clinically evaluated by SLPs at the point of recruitment hence it could not be determined if the recommendations provided by the LSEds were accurate reflections of the LS participants' language developmental profiles. Secondly, the three LS participants who were identified to require a second package of language learning support services and further evaluation were not followed up over time. Therefore, it was unknown if these three LS participants continued to demonstrate language concerns over time. Similarly, for the six LS participants who were discharged from DSP services, it was unknown if any of them continued to demonstrate language concerns and were referred back to the DSP.

Nonetheless, the above findings with this small group of LS participants, combined with the high classification accuracy of SLT and TD participants into their respective groups (Section 10.2.4.3), have provided some promising evidence under *Extrapolation*: that the construct of language learning potential assessed on the DA predicted scores that correspond to the bilingual children's language developmental profiles (Table 10-5). In other words, when compared with the CELF P2 EV subtest, the DA of word learning skills was less biased in evaluating the language skills of bilingual children and identified those at risk of LI more accurately. Commercially-available standardised vocabulary measures (i.e. CELF P2 EV subtest) are mostly designed for use with monolingual children and are not appropriate for evaluating the language skills of bilingual children and therefore identifying those at risk of LI (e.g. De Lamo White & Jin, 2011). In fact, the use of commercially-available standardised assessments may lead to an over-referral of bilingual children to speech and language therapy when they, in fact, have normal language learning abilities.

Table 10-5. Evidence to support the claim under Extrapolation.

Inference	Claim	Assumptions	Evidence
Extrapolation	The construct of language learning potential as assessed on the DA predicts scores that correspond to bilingual children's language developmental profiles.	The scores obtained on the DA can accurately predict bilingual children's language developmental profiles.	Scores on the DA correctly classified 93.9% bilingual children identified for on-going speech and language therapy (SLT participants) and typically developing bilingual preschool children (TD participants) into their respective groups. (Specificity: 95.8%; Sensitivity: 88.9%). Nine children with language concerns (LS participants) were assessed on the DA. Eight out of nine of them obtained scores on the DA that corresponded to the recommendations from their end-of-intervention reports.

10.2.5 Evaluation of the Validity Argument Framework

The use of the Validity Argument Framework (Kane, 2006; 2013) to gather evidence to support the interpretation and use of a DA process is, to the knowledge of the researcher, the first in the field of language testing. Other DA studies (e.g. Bain & Olswang, 1995 Camilleri & Botting, 2013) have previously evaluated and established validity evidence based on the psychometric properties of the DA and not on the validity of the interpretation and use of testing outcomes of the DA in evaluating the language skills of bilingual children, and therefore identifying those at risk of LI.

Findings from this study provide some preliminary evidence gathered using the Validity Argument Framework (Kane, 2006; 2013) to support the use and interpretation of the DA of word learning skills to validly: 1) differentiate typically developing Singaporean bilingual preschool children and Singaporean bilingual preschool children referred for on-going speech and language therapy and; 2) identify or screen for Singaporean bilingual preschool children

at risk of LI. In other words, the DA of word learning skills has potential as a less biased alternative assessment approach to evaluate the language skills of bilingual children and therefore identify those at risk of LI regardless of their prior language experience. Additionally, findings from this study found that the CELF P2 EV subtest, a commonly used standardised assessment, cannot validly evaluate the language skills of Singaporean bilingual children and therefore identify those at risk of LI. The performance of Singaporean bilingual children on the CELF P2 EV subtest can be influenced by their past language experience. Results from the current study found that MD participants, regardless of their language developmental profiles, always performed significantly more poorly than their ED counterparts on the CELF P2 EV subtest. Moreover, 37.9% of TD MD participants obtained a scaled score of 6 and below which indicated that they had poor expressive vocabulary skills and were at risk of LI when they, in fact, were screened on the SEAPT to be not at risk of LI prior to inclusion into the TD group.

During the development of the DA of word learning skills, the chain of evidence gathered under the Validity Argument Framework (Kane 2006, 2013) was of confirmationist bias. This was intentional so as to advocate for the proposed use and interpretation of DA as a less biased assessment approach for assessing and identifying LI among bilingual children with variable language experience.

10.3 Improving the Clinical Acceptability of DA in Bilingual Clinical Settings

Assessing bilingual children's existing language knowledge on static assessments, such as standardised assessments, works on the assumption that all children have similar language experience and are exposed to similar language concepts and vocabulary (De Lamo White & Jin, 2011). However, this is often not the case for bilingual children as even language exposure and experience are variable among bilingual children from the same community who are exposed to the same two languages. Research has repeatedly found that bilingual children's performance on standardised assessment is dependent on the amount of input that they have

had in the language that is being assessed (e.g. Brebner et al., 2016; Hoff & Core, 2013; Thordardottir, 2011; Teoh et al., 2012).

On the other hand, assessing bilingual children's language learning potential using DA can minimise the test bias associated with previous language experience and exposure. Instead of evaluating existing language knowledge, DA evaluates children's ability/potential to learn new language materials. Based on the processing-based account of LI, limitations in cognitive processing capacity have resulted in children with LI demonstrating deficits in encoding and rehearsing new language materials in their working memory (Bedore & Peña, 2008; Leonard et al., 2007; Paradis, 2010). Therefore, children with LI often present with a slower rate of learning, often requiring twice the amount of time and input to learn new language materials (i.e. novel words) (Gray 2003). Typically developing children with intact cognitive processes are able to encode and rehearse new language information at a much faster rate. Therefore, evaluating bilingual children's potential to learn new language material (i.e. novel words) in a DA process can reveal their actual language learning potential and assist SLPs in identifying bilingual children with LI with accuracy.

There is an emerging body of research on DA and its use to discriminate bilingual children with and without LI in vocabulary and word learning studies (Camilleri & Botting, 2013; Kapantzoglou et al., 2012; Peña et al., 2001), narrative studies (Peña et al., 2006) and sentence structure studies (Hasson et al, 2012). However, despite favourable results from the literature, many SLPs working in bilingual clinical settings are still not using DA as a choice assessment (Caesar & Kohler, 2007; Teoh et al., 2017). Reasons for the lack of adoption of DA in bilingual clinical settings were identified in the literature review (Chapter 4, Section 4.4). The lack of a structured protocol to deliver mediated teaching instructions and a lack of an objective framework in scoring children's modifiability may have hindered SLPs from readily adopting DA approaches into their practice. In addition, there is also a lack of exploratory research on

DA approaches for use in predominantly bilingual countries in discriminating the language learning potential of bilingual children with and without LI.

In previous studies, one of the major limitations in translating DA from research to clinical utility was that DA often consisted of testing and teaching sessions that spanned over multiple sessions (e.g. Peña et al., 2001; 2006; Ukrainetz et al., 2000). Therefore, it is not surprising that SLPs prefer to use standardised assessments that can be administered and scored in a single session even though they are known to be biased when evaluating bilingual children's language ability. It is only in recent years that DA studies have attempted to simplify and shorten DA into a single session (e.g. Kapantzoglou et al., 2012; Petersen et al., 2017). However, in these studies, the evaluation of children's modifiability was still based using examiners' subjective ratings of the children's responsiveness to teaching instructions on modifiability rating scales/checklists. The subjective aspect of modifiability ratings is likely to hinder the clinical acceptability of DA as SLPs have to undergo additional professional training to be familiar with the allocation of modifiability ratings accurately.

During the development of the DA of word learning skills in the current study, the researcher took into consideration test design factors that could increase the clinical acceptability of DA. The researcher also considered SLPs' responses regarding their assessment practices (Chapter 3, Section 3.4.4) in which SLPs identified 'time required to administer the assessment' and the 'ease of scoring the assessment' as important factors to consider when developing a new assessment measurement for their local population.

To recap, the DA of word learning skills was designed to be:

1. Administered and scored in a single session of 30 -45 minutes.

- A brief administration and scoring process would increase the clinical acceptability of DA as a quick screening alternative assessment measure for identifying bilingual children at risk of LI.

2. Administered in a 'Test-Teach-Retest' framework using a structured format and scripted teaching instructions that incorporated both MLE and GP principles.

    - A structured administration protocol would allow examiners to administer the DA of word learning skills without having to undergo any additional training on the provision of MLE or GP teaching instructions. This would contribute to the clinical acceptability of DA as an alternative assessment measure.

3. Children's performance was scored objectively based on: 1) the number of mediation prompts provided by the examiner in the 'Teach' phase to identify and name all four target words, and 2) the number of target words that were successfully identified and named in the 'Retest' phase.
    - A structured scoring framework would allow examiners to objectively analyse and interpret the children's performance on the DA of word learning skills, increasing the reliability of their observations. This would contribute to the clinical acceptability of DA as an alternative assessment measure.

To the researcher's knowledge, this study is the first in the field of language testing to explore the use of DA in a predominantly bilingual country (i.e. Singapore) to discriminate the language learning potential between typically developing bilingual children and bilingual children identified to require on-going speech and language therapy. The findings of the study as discussed have provided preliminary validity evidence that the DA of word learning skills

can validly evaluate the language learning potential of bilingual children, and therefore identify those at risk of LI. More specifically, compared to a commercially-available standardised measure, the DA of word learning skills demonstrated higher classification accuracy in discriminating between typically developing bilingual children and bilingual children identified to require on-going speech and language therapy. It is hoped that the findings of this study can increase the awareness that DA is a viable alternative assessment approach to be included as part of as assessment battery for assessing the language skills of bilingual children in predominantly bilingual countries.

10.4 Limitations and Future Research

The findings from the study indicated that DA compared to commercially-available standardised assessment is a more accurate assessment measure for screening and validly identifying bilingual children at risk of LI. However, the validity evidence gathered to support the DA of word learning skills for differentiating bilingual children with and without LI and identifying bilingual children with LI in this study is exploratory. The DA of word learning skills will need further methodological reviews and pilot testing before it can be developed as an actual screening assessment measure to validly identify bilingual children at risk of LI.

Furthermore, validation is not an end-point but an on-going process (Cook et al., 2015). Although the scoring of children's performance on the DA of word learning skills is developed to be objective, it is unknown if SLPs would find the scoring protocol of the DA of word learning skills to be clinically-user friendly. It is also unknown if the objective scoring protocol designed allowed scoring to be reliable across examiners. Therefore, it will be useful for inter-rater reliability for the DA of word learning skills to be established in future studies. This would add to the validity evidence under *Scoring* inference of the Validity Argument Framework (Kane, 2006; 2013).

Next, the study was only conducted with Singaporean English-Mandarin bilingual preschool children, and broader testing within the Singaporean bilingual preschool community (i.e. English-Malay bilingual children, English-Tamil bilingual children) is required. This is to explore whether the DA of word learning skills can accurately differentiate between bilingual children with and without LI from different ethnic groups who do not speak the same language other than English. At the same time, the DA of word learning skills may be preliminarily explored for use with other bilingual communities, especially in predominantly bilingual countries where there is still a lack of appropriate local language assessment measures. This would add to the validity evidence under *Generalisation* inference of the Validity Argument Framework (Kane, 2006; 2013).

Last but not least, for the DA of word learning skills to be used as a potential screening measure to identify bilingual children at risk of LI, classification cut-off points are required. Without a standard for what is likely impaired and what is an average performance, SLPs may not be confident in using DA in making clinical decisions (Petersen et al., 2017). Therefore for future studies, it may be helpful to determine cut-off points and administer the DA of word learning skills to a cohort of bilingual children while following up on their progress over time. This is to determine whether children who are identified as at risk of LI based on the cut-off points continued to present with language difficulties in the long term. This would ideally involve a study where the cohort of bilingual children are evaluated by SLPs at 6 month and 12 month intervals, in addition to teachers' reports of these children's language and learning ability in class at these time intervals. This would add to the validity evidence under *Extrapolation* inference of the Validity Argument Framework (Kane, 2006; 2013).

10.5 Chapter Summary

The Validity Argument Framework (Kane 2006; 2013) was adopted in Phase Two of the study to present and discuss the findings on whether the DA of word learning skills compared to

standardised assessment of vocabulary can validly discriminate the language learning potential of typically developing Singaporean bilingual preschool children and Singaporean bilingual preschool children who were identified for on-going speech and language therapy.

Evidence was gathered for the claims under each inference of the Validity Argument Framework (Kane 2006; 2013). A discussion on how the DA process developed in Phase Two in the study attempted to considered test-design elements to increase the clinical acceptability of DA was also presented. Finally, the limitations and directions for future research were presented.

Chapter 11 Overall Conclusion

11.1 Introduction

In this chapter, a recap of the aims of the thesis is first presented. Next, the findings from Phase One and Phase Two of the study are summarised. This is followed by a discussion of the clinical implications that can be inferred from the findings. Lastly, a concluding statement summarising the intended outcomes of the study will be presented.

11.2 Summary of Thesis Aims

The valid assessment and identification of LI among young bilingual children is challenging. Young bilingual children, even from the same bilingual community who speak the same two languages, are known to have variable language experiences in each language they are exposed to (Bedore & Peña, 2008; Kohnert, 2009; Paradis, 2010). Assessment bias may result if a bilingual child's language skill is only assessed based on his or her performance on commercially-available standardised assessments where only existing language knowledge in a single language is evaluated (De Lamo White & Jin, 2011). Moreover, most commercially-available standardised assessments are created and normed for use with monolingual-speaking children of the language (De Lamo White & Jin, 2011). For these reasons, commercially-available standardised assessments should not be used to evaluate bilingual children's language skills as they cannot be used to validly screen or identify LI among bilingual children.

An emerging body of research has now instead focused on adopting alternative assessment approaches for evaluating the language skills of bilingual children instead (De Lamo White & Jin, 2011; Hemsley et al., 2014; Laing & Kamhi, 2003; Pieretti & Roseberry-McKibbin, 2016). However, the research on recommended bilingual assessment practices and exploration of alternative assessment approaches has largely focused on minority bilingual communities residing in predominantly monolingual countries (e.g. UK, USA). On the other hand, the

research on recommended bilingual assessment practices and exploration of the feasibility of alternative assessment approaches for use with bilingual children residing in predominantly bilingual countries has been sparse.

Little is known about the assessment practices and the assessment challenges faced by SLPs working in predominantly bilingual countries and whether SLPs working in these countries are adopting recommended alternative assessment approaches. The main aim of this study was therefore to address the gap in the research on the use of alternative assessment approaches in evaluating the language skills of bilingual children from predominantly bilingual countries. This was to contribute to the existing literature on the use of appropriate assessment approaches in evaluating the language skills of young bilingual children with variable language experience and identifying those at risk of LI.

This thesis addressed this gap in the research in two phases. In Phase One, a survey study was conducted to understand the assessment practices and challenges faced by SLPs working in Singapore, an English-speaking and predominantly bilingual country in South-East Asia. One of the main findings found that there was fair to limited use of alternative assessment approaches. This could be attributed to the fact that no local study has explored the use of alternative assessment approaches with the Singaporean bilingual population.

To increase the adoption of alternative assessment approaches to evaluate the language skills of bilingual children in predominantly bilingual countries like Singapore, Phase Two aimed to develop and explore the use of a DA process. Specifically, Phase Two aimed to gather validity evidence regarding whether a DA process can validly differentiate the language learning potential between bilingual children with and without LI, and therefore accurately identifying those at risk of LI. The intention was not to develop a finalised DA tool but to gather evidence for the proposed interpretation and use of the DA process.

11.3 Summary of Thesis Findings

11.3.1 Phase One: Survey study

To understand the assessment practices and challenges faced by SLPs working in a predominantly bilingual country, a survey study was conducted in Singapore. A major finding from the survey study was that SLPs in Singapore were choosing commercially-available standardised assessments over alternative assessment approaches when assessing the language skills of Singaporean bilingual children. This is despite the fact that these commercially-available standardised assessments that are designed and normed for use with monolingual English-speaking children may not validly evaluate the language skills of Singaporean bilingual children. The over-reliance on commercially-available standardised assessments was deduced to be attributed to:

1. Little or no information on local normative and language developmental data for the languages spoken in Singapore. This likely has resulted in SLPs assuming that the developmental trajectories of Singaporean bilingual children in each language they speak are comparable to monolingual norms.
2. The lack of local practice guidelines on recommended bilingual assessment practices and alternative assessment approaches. This may have resulted in SLPs adopting assessment approaches that are used to assess the language skills of monolingual children.
3. The lack of local studies in exploring the feasibility and utility of alternative assessment approaches in evaluating the language skills of bilingual children and accuracy in identifying LI among bilingual children. Therefore, even if SLPs were aware of recommended alternative assessment approaches, they were likely not confident in adopting alternative assessment approaches in their practice.

11.3.2 Phase Two: Development of the DA Process and Gathering of Validity Evidence

The lack of exploration of the feasibility of alternative assessment for use with bilingual children residing in predominantly bilingual countries combined with the findings from Phase One of the study led to the initiation of Phase Two of the study. Phase Two aimed to explored whether DA, an alternative assessment approach, when compared to a commercially available standardised assessment, could validly differentiate between Singaporean bilingual preschool children with and without LI, and therefore identify Singaporean bilingual preschool children at risk of LI. Due to absence of local data on the developmental trajectories of the language spoken in Singapore, DA was chosen over other alternative assessment measures (i.e. language sampling, criterion-referenced measures) as it does not require local normative data for comparison. Instead of evaluating existing language knowledge like most other assessment approaches (i.e. standardised assessments, language sampling), DA assesses children's language ability by evaluating their potential to learn new language materials.

To demonstrate that DA can validly assess the language skills of bilingual children, validity evidence was gathered systematically using the Validity Argument Framework (Kane 2003, 2006). Using the framework, a chain of validity evidence based from the development of the DA task to decisions based on the scores on the interpretation and use of testing outcomes of the DA was gathered.

The DA process was first developed by reviewing the literature to identify a suitable DA language task. Due to the variable language experiences that bilingual children present with, it is least biased to evaluate their ability to learn new language materials under the processing-based account. Word learning, a processing based task, was selected as poor word learning performance was identified in the literature review as a potential behavioural indicator of LI among young bilingual children. The evidence gathered supported the first claim made under the inference *Domain Definition*: that is observations of performance on the DA reveal a

bilingual children's language learning potential, and this can be used to determine whether a bilingual child is at risk of LI.

Next, evidence was gathered to support the application and accuracy of the scoring rubric used to evaluate children's language learning potential on the DA task. A 'Test-Teach-Retest' format was chosen. During the development of the DA of word learning skills, there was careful selection of the stimuli (i.e. four non-words with high phonotactic probability and low neighbourhood density) so as to maximise the difference in word learning abilities between children with and without LI. In addition, the administrative process was designed to present each target word only nine times to ensure that sufficient opportunities were provided for word learning to take place among typically developing children while the task remained challenging for children with LI. Next, to ensure that the rubric for scoring the children's responses on the DA task corresponded to their language learning performance, an objective scoring framework that incorporated both GP and MLE principles was developed to score the children's modifiability in the 'Teach' phase (i.e. scored based on the number of mediation prompts provided to identify and name all four target words) and ability to recall and identify to name the target words in the 'Retest' phase (i.e. scored based on the number of target words identified and named independently). A pilot study was subsequently conducted to ensure that the administrative procedure and scoring rubric were appropriate. The evidence gathered supported the second claim made under the inference *Scoring:* that is observations of performance on the DA task are evaluated to provide observed scores reflective of language learning potential.

In the third step, evidence was gathered to demonstrate that the DA of word learning skills was able to capture accurate, consistent and relevant summaries of bilingual children's language learning abilities. The evidence was gathered by comparing the performances of two groups of bilingual preschool children on the DA of word learning task: children who required on-going

speech and language therapy (SLT) and typically developing children (TD). Results showed that there were significant differences in performance between SLT and TD group in all four scoring aspects of the DA. SLT children required more mediation prompts than TD children to identify and name all four target words in the 'Teach' phase. Despite receiving more mediation prompts, SLT children identified and named less target words compared to TD children in the 'Retest' phase. In addition, within each group, there was no difference in the scores obtained between children with variable language experiences (i.e. TD English dominant vs TD Mandarin dominant). This demonstrated that the DA of word learning skills was able to capture the children's language learning abilities with minimal bias as their performance was not influenced by their prior language experience. This was unlike the results on the CELF P2 EV subtest, a standardised vocabulary assessment, where their performance was influenced by their prior language experience. The evidence gathered overall supported the third claim made under the inference *Generalisation*: that is the observed scores on the DA reflect language learning potential that are consistent among bilingual preschool children with the same language developmental profiles regardless of their prior language experience.

Finally, evidence was gathered to demonstrate that the scores obtained on DA of word learning skills could predict bilingual children's language developmental profiles in the 'real' world. The evidence was gathered by running a logistic model to evaluate how accurately both groups of children's DA scores predicted their group membership (i.e. SLT or TD). Results found that the model classified 93.9% of the children accurately, with 88.9% sensitivity and 95.8% specificity. This contrasts with the CELF P2 EV subtest, in which only 74.2% of the children were classified accurately with 71.1% specificity and 66.7% sensitivity. A third group of children with language concerns (LS participants) was evaluated on the DA of word learning skills and CELF EV subtest; their performances in both assessments were correlated with the recommendations from LSEds in their end-of-intervention reports. Based on the scores

obtained on the DA of word learning skills, eight out of nine children had group predictions that corresponded to the recommendations in their end-of-intervention reports. However, on the CELF P2 EV subtest, only four out of nine children obtained scores that corresponded to the recommendations in their end-of-intervention reports. The findings overall demonstrated that compared to a standardised assessment of vocabulary, the DA of word learning skills was able to more accurately classify bilingual children with and without LI. Therefore, the evidence gathered from both sets of results supported the fourth claim made under the inference *Extrapolation*: that is the construct of language learning potential as assessed on the DA predicts a score that corresponds to the child's language developmental profile in the real world.

The overall findings from Phase Two provided validity evidence that the DA of word learning skills can validly evaluate the language learning potential of bilingual children, and therefore identify those at risk of LI. More specifically, compared to a commercially-available standardised measure, the DA of word learning skills demonstrated higher classification accuracy in discriminating between typically developing bilingual children and bilingual children identified to require on-going speech and language therapy.

11.4 Clinical Implications

11.4.1 Increasing the number bilingual SLPs does not resolve the assessment challenges faced in assessing bilingual children

One of the main findings from Phase One (survey study) was that although most SLPs working in a predominantly bilingual country like Singapore were bilinguals, they did not always speak the same languages as the bilingual children on their caseloads. In fact, bilingual children on their caseloads often spoke a diverse range of languages. In other words, due to the diverse range of languages spoken in a predominantly bilingual country or community, mismatch of languages spoken by the SLP and children on their caseload will still occur. Even in cases where the languages of the SLP and the children match, SLPs still need culturally and

linguistically appropriate assessment tools to be able to assess the children's language skills. Unfortunately, culturally and linguistically appropriate assessments and local data on the developmental trajectories of language spoken in bilingual communities are rarely available. Therefore, the solution of increasing the number of bilingual SLPs suggested by some researchers (D'Souza et al., 2012; Perrie & Core, 2006) to meet the demands of providing speech and language therapy services to bilingual children is not a practical solution to all of the issues.

11.4.2 Commercially-available standardised assessments should not be used as a stand-alone assessment to evaluate the language skills of bilingual children

With regard to the lack of appropriate local assessment tools for bilingual populations and understanding of the developmental trajectories of the local languages spoken within each bilingual community, this study can also offer insights for bilingual assessment practices around the world.

The findings from Phase One and others (e.g. Caesar & Kohler, 2007; McLeod, 2014; Williams & McLeod, 2012) show there is still an over-reliance on commercially-available standardised assessments that were designed and normed for use with predominantly monolingual English-speaking children when assessing the language skills of bilingual children. Solutions such as assessment modification (i.e. contents to be culturally and linguistically appropriate) and 're-standardisation' have been explored but may not be most appropriate. This is because bilingual populations throughout the world are often heterogeneous in nature, even within a bilingual community. As Gathercole (2014: p.360, 359) has aptly described, 'variation is the norm, not the exception' and 'one size does not fit all'. Bilingual children's abilities in each language are known to evolve over time as a result of changing language experiences and individual differences.

If commercially-available standardised assessments are still preferred, SLPs must be aware of their limitations as they were not designed for use to assess the language skills of bilingual children. The findings from Phase Two and previous studies (Camilleri & Law, 2007; Teoh et al., 2012) reiterate the fact that commercially-available standardised assessments do not fairly and validly assess the language skills of bilingual children, especially if the bilingual child's dominant language (i.e. Non-English dominant) does not match to that of the language the assessment (i.e. English) was designed to assess. This may lead to an over-identification of bilingual children for speech and language therapy services when they, in fact, have normal language learning abilities. At the same time, it may also lead to an under-identification of LI if SLPs attribute bilingual children's poor performance on the standardised assessment as language difference.

If commercially-available standardised assessments have to be used, it should be used as part of an assessment battery where alternative assessment approaches are also used to gather information about the bilingual child's language ability. Bilingual children's performance on commercially-available standardised assessments should only be used to provide qualitative descriptions of a bilingual child's language ability (De Lamo White & Jin, 2011). Their performance cannot be meaningfully scored for comparison with normative data for monolingual-speaking children (De Lamo White & Jin, 2011). At the same time, information on language exposure, language ability, and developmental history through parent report/questionnaire should be gathered as it is known to increase accurate identification of LI among bilingual children (Paradis & Duncan, 2010; Paradis, 2016).

11.4.3 DA should be part of the assessment battery when assessing bilingual children

In the absence of local normative data and developmental trajectories for the languages spoken in bilingual community, DA is a viable assessment approach. Findings from Phase Two of the study showed that DA – an alternative assessment approach, can accurately differentiate

bilingual children with and without LI from variable language backgrounds and consequently identify bilingual children at risk of LI. This finding adds to the emerging research (i.e. De Lamo White & Jin, 2011; Hemsley et al., 2014; Pieretti & Roseberry-McKibbin, 2016) on the recommended use of DA to assess the language learning potential of bilingual children and therefore identify those at risk of LI. Despite the evidence, DA is still not a frequently used alternative assessment measure. It is likely that clearer markers for modifiability (i.e. language learning potential) need to be identify in future research. Future research is also needed to demonstrate that poor language learning potential in a DA context can be a definite indicator for language impairment among both monolingual and bilingual children.

To increase the adoption of alternative assessment approaches such as DA, especially in bilingual clinical settings, there is also a need to explore and develop alternative assessment approaches for use in local bilingual communities (i.e. language sampling, DA). This is to increase the clinical awareness and acceptability of alternative assessment approaches.

11.4.4 Recommended practices on bilingual assessment practices are needed

Guidelines or/and position papers on recommended bilingual assessment practices and intervention in countries (i.e. Singapore) when bilingualism is the norm should be developed. This may reduce the over-adoption of monolingual assessment practices and create increased awareness of recommended bilingual assessment practices. In the meantime, SLPs can refer to literature on recommended bilingual assessment practices and alternative assessment approaches (i.e. Bedore & Peña, 2008; De Lamo White & Jin, 2011; Kohnert, 2010; Paradis, 2016; Pieretti & Roseberry-McKibbin, 2016).

11.5 Concluding Statement

The findings of this study addressed the gap in the research on the use of alternative assessment approaches in evaluating the language skills of bilingual children from predominantly bilingual countries. Specifically, the findings gathered in Phase Two of the study gathered validity

evidence and demonstrated that DA compared to commonly used commercially-available standardised assessments can evaluate the language skills of bilingual children with variable language experience with less bias.

It is hoped that that the findings from the study will not only increase the awareness of recommended alternative assessment practices with bilingual children in predominantly bilingual countries where research is sparse but also pave the way for the clinical acceptability and further development of DA as an appropriate assessment approach for evaluating the language skills of bilingual children and to identify those at risk of LI. Specifically, in Singapore, it is hoped that the evidence gathered can pave the way for clinical acceptability and future development of a localised DA tool.

References

Akbari, R. (2012). Validity in language testing. In C. A. Coombe (Ed.), *The Cambridge guide to second language assessment* (pp. 30-36). New York: Cambridge University Press.

Allied Health Professions Council Singapore. (2014). *Report on the registration of occupational therapists, physiotherapists, speech-language therapists*. Allied Health Professions Council, Ministry of Health, Singapore.

Alt, M. (2011). Phonological working memory impairments in children with specific language impairment: Where does the problem lie? *Journal of Communication Disorders, 44*(2), 173-185. doi:10.1016/j.jcomdis.2010.09.003

Alt, M., & Plante, E. (2006). Factors that influence lexical and semantic fast mapping of young children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 49*(5), 941-954. doi:10.1044/1092-4388(2006/068)

Alt, M., Plante, E., & Creusere, M. (2004). Semantic features in fast- mapping: Performance of preschoolers with specific language impairment versus preschoolers with normal language. *Journal of Speech, Language, and Hearing Research, 47*(2), 407.

Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research, 52*(4), 899-914. doi:10.1044/1092-4388(2009/08-0099)

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189-208. doi:10.1016/S0021-9924(03)00019-4

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review, 105*(1), 158-173. doi:10.1037/0033-295X.105.1.158

Bain, B. A., & Olswang, L. B. (1995). Examining readiness for learning two-word utterances by children with specific expressive language impairment: Dynamic assessment validation. *American Journal of Speech-Language Pathology, 4*(1), 81-91. doi:10.1044/1058-0360.0401.81

Bedore, L., & Peña, E. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism, 11*(1), 1-29. doi:10.2167/beb392.0

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*(5), 1185. doi:10.1044/jslhr.4105.1185

Bedore, L. M., & Leonard, L. B. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(4), 905-924. doi:10.1044/1092-4388(2001/072)

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language & Communication Disorders, 49*(4), 381-415. doi:10.1111/1460-6984.12101

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry, 37*(4), 391-403. doi:10.1111/j.1469-7610.1996.tb01420.x

Bloom, L. (2000). The intentionality model of word learning. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. L. Woodward, N. Akhtar, M. Tomasello, & G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 19 - 50). Oxford: Oxford University Press.

Bortolini, U., Caselli, M. C., & Leonard, L. B. (1997). Grammatical deficits in Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*(4), 809-820. doi:10.1044/jslhr.4004.809

Botting, N., & Conti-Ramsden, G. (2001). Non- word repetition and language development in children with specific language impairment (SLI). *International Journal of Language and Communication Disorders, 36*(4), 421-432. doi:10.1080/13682820110074971

Bowen, C. (1996). *The quick screener*. Retrieved from www.speech-language-therapy.com

Brebner, C. (2002). *The Singapore English action picture test*. Published with permission of Speechmark Publishing Ltd.

Brebner, C. (2010). The development of the Singapore English action picture test: An expressive language screening assessment for Chinese Singaporean preschool children. In M. Cruz-ferreira (Ed.), *Multilingual norms* (pp. 323-341). Bern: Peter Lang Publishing Group.

Brebner, C., McCormack, P., & Liow, S. R. (2016). Marking of verb tense in the English of preschool English-mandarin bilingual children: Evidence from language development profiles within subgroups on the Singapore English action picture test. *International Journal of Language & Communication Disorders, 51*(1), 31-43. doi: 10.1111/1460-6984.12181

Burton, V. J., & Watkins, R. V. (2007). Measuring word learning: Dynamic versus static assessment of kindergarten vocabulary. *Journal of Communication Disorders, 40*(5), 335-356. doi:10.1016/j.jcomdis.2006.06.015

Caesar, L. G., & Kohler, P. D. (2007). The state of school- based bilingual assessment: Actual practice versus recommended guidelines. *Language, speech, and hearing services in schools, 38*(3), 190-200. doi:10.1044/0161-1461(2007/020)

Camilleri, B., & Botting, N. (2013). Beyond static assessment of children receptive vocabulary: The dynamic assessment of word learning (DAWL). *International Journal of Language and Communication Disorders, 48*(5), 565-581. doi:10.1111/1460-6984.12033

Camilleri, B., & Law, J. (2007). Assessing children referred to speech and language therapy: Static and dynamic assessment of receptive vocabulary. *Advances in Speech Language Pathology, 9*(4), 312-322. doi:10.1080/14417040701624474

Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievements. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82-115). New York, NY: Guildford Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17-29.

Carlson, J., & Wiedl, K. H. (2000). The validity of dynamic assessment. In C. L. J. G. Elliott (Ed.), *Dynamic assessment: Prevailing models and applications* (pp. 681-712). New York, NY: Elsevier Science.

Carrow-Woolfolk, E. (1999). *Comprehensive assessment of spoken language.* Bloomington, MN: Pearson Assessments.

Carson, K., Gillon, G., & Boustead, T. (2011). Computer-administrated versus paper-based assessment of school-entry phonological awareness ability. *Asia Pacific Journal of Speech, Language and Hearing, 14*(2), 85-101. doi:10.1179/136132811805334876

Chang, C. S. E., Young, S. E.-L., Rickard-Liow, S., & Yin Chong, C. X. (2015). Story-retelling abilities of bilingual children with non-syndromic cleft lip and/or palate. *Speech, Language and Hearing, 18*(2), 91-101. doi:10.1179/2050572814Y.0000000054

Chapelle, C. (2008). The tofel validity argument. In J. Jamieson, M. Enright, & C. Chapelle (Eds.), *Building a validity argument for the test of English as a foreign language* (pp. 317-352). London, England: Routledge.

Chapelle, C. A., Enright, M. K., & Jamieson, J. (2010). Does an argument-based approach to validity make a difference? *Educational Measurement: Issues and Practice, 29*(1), 3-13. doi:10.1111/j.1745-3992.2009.00165.x

Chapelle, C. A., & Voss, E. (2014). Evaluation of language tests through validation research. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1081-1097). Chichester, United Kingdom: John Wiley and Sons Ltd.

Chomsky, N. (1957). *Syntactic structure*. The Hauge: Mouton.

Chua, S. K. C. (2011). Singapore's e(si)nglish-knowing bilingualism. *Current Issues in Language Planning, 12*(2), 125-145. doi:10.1080/14664208.2011.602816

Clahsen, H. (1989). The grammatical characterization of developmental dysphasia. *Linguistics, 27*(5), 897. doi:10.1515/ling.1989.27.5.897

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Cleave, P. L., Girolametto, L. E., Chen, X., & Johnson, C. J. (2010). Narrative abilities in monolingual and dual language learning children with specific language impairment. *Journal of Communication Disorders, 43*(6), 511-522. doi:10.1016/j.jcomdis.2010.05.005

Cohen, J. (1969). *Statistical power analysis for the behavioural sciences*. New York: New York, Academic Press.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42*(6), 741-748. doi:10.1017/S0021963001007600

Conti-Ramsden, G., & Durkin, K. (2011). Specific language impairment *Child psychology and psychiatry: Frameworks for practice* (pp. 180-185): John Wiley & Sons, Ltd.

Conti-Ramsden, G., & Windfuhr, K. (2002). Productivity with word order and morphology: A comparative look at children with SLI and children with normal language abilities. *International Journal of Language & Communication Disorders, 37*(1), 17-30. doi:10.1080/13682820110089380

Conti-Ramsden, G., Mok, P. L. H., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (sli): Strengths and difficulties in social, emotional and behavioral functioning. *Research in developmental disabilities, 34*(11), 4161-4169. doi:https://doi.org/10.1016/j.ridd.2013.08.043

Cook, D. A., Brydges, R., Ginsburg, S., & Hatala, R. (2015). A contemporary approach to validity arguments: A practical guide to kane's framework. *Medical Education, 49*(6), 560-575. doi:10.1111/medu.12678

Creswell, J. W. (2014). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Harlow, Essex: Pearson.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. doi:10.1037/h0040957

Cruz-Ferreira, M., & Ng, B. C. (2010). Assessing multilingual children in multilingual clinics: Insights from Singapore. In M. Cruz-Ferreira (Ed.), *Multilingual norms* (pp. 343-396). Frankfurt: Peter Lang Publishing Group.

Davies, A., & Elder, C. (2005). Validity and validation in language testing. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 795-813). Mahwah: Taylor and Francis.

De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language and Communication Disorders, 46*(6), 613-627. doi:10.1111/j.1460-6984.2011.00049.x

DiStefano, C., & Morgan, G. (2011). Examining classification criteria: A comparison of three cut score methods. *Psychological Assessment, 23*(2), 354-363. doi:10.1037/a0021745

Dixon, L. Q., Wu, S., & Daraghmeh, A. (2012). Profiles in bilingualism: Factors influencing kindergartners' language proficiency. *Early Childhood Education Journal, 40*(1), 25-34. doi:10.1007/s10643-011-0491-8

Dixon, W. E., Salley, B. J., & Clements, A. D. (2006). Temperament, distraction, and learning in toddlerhood. *Infant behavior & development, 29*(3), 342-357. doi:10.1016/j.infbeh.2006.01.002

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136-1146. doi:10.1044/jslhr.4105.1136

D'Souza, C., Bird, E. K.-R., & Deacon, H. (2012). Survey of Canadian speech-language pathology service delivery to linguistically diverse clients. *Canadian Journal of Speech-Language Pathology and Audiology, 36*(1), 18-39.

Duinmeijer, I., De Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language & Communication Disorders, 47*(5), 542-555. doi:10.1111/j.1460-6984.2012.00164.x

Dunn, L. H., Dunn, D. M., Whetton, C., & Burley, J. (1997). *The British picture vocabulary scale II- 2nd edition*. Berkshire: NFER-Nelson.

Earle, F. S., Gallinat, E. L., Grela, B. G., Lehto, A., & Spaulding, T. J. (2017). Empirical implications of matching children with specific language impairment to children with

typical development on nonverbal IQ. *Journal of Learning Disabilities, 50*(3), 252-260. doi:doi:10.1177/0022219415617165

Ebert, K. D., & Kohnert, K. (2016). Language learning impairment in sequential bilingual children. *Language Teaching, 49*(3), 301-338. doi:10.1017/S0261444816000070

Ehrler, D. J., & McChee, R. L. (2008). *Primary test of non-verbal intelligence* Austin, Texas: PRO-ED, Inc.

Elli Weismer, S., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*(5), 1249-1260. doi:10.1044/jslhr.4205.1249

Elli Weismer, S., & Evans, J. L. (2002). The role of processing limitations in early identification of specific language impairment. *Topics in Language Disorders, 22*(3), 15-29.

Feuerstein, R., Rand, Y., & Hoffman, M. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. Baltimore, MD: University Park Press.

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics* (4th edition. ed.): Los Angeles SAGE Publications Ltd.

Fung, R. S.-Y. (2008). Characteristics of Chinese in relation to language disorders. In S. P. Law, A. M.-Y. Wong, & B. Weekes (Eds.), *Language disorders in speakers of Chinese*: Multilingual Matters.

Gallinat, E., & Spaulding, T. J. (2014). Differences in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: A meta- analysis. *Journal of Speech, Language, and Hearing Research, 57*(4), 1363-1382. doi:10.1044/2014_JSLHR-L-12-0363

Gardner, M. F. (1990). *Expressive one-word picture vocabulary test*. Novato,CA: Acaedmic Therapy Publications.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45*(3), 468-492. doi:https://doi.org/10.1006/jmla.2000.2784

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*(3), 336-360. doi: 10.1016/0749-596X(90)90004-J

Gathercole, V. C. M. (2014). Bilingualism matters: One size does not fit all. *International Journal of Behavioral Development, 38*(4), 359-366. doi:10.1177/0165025414531676

Gillam, R. B., Cowan, N., & Marler, J. A. (1998). Information processing by school-age children with specific language impairment: Evidence from a modality effect paradigm. *Journal of Speech, Language, and Hearing Research, 41*(4), 913-926. doi:10.1044/jslhr.4104.913

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews, 11*(3), 173-179. doi:10.1002/mrdd.20068

Gn, L. W. E., Brebner, C., & McCormack, P. (2014). A preliminary report on the English phonology of typically developing English– mandarin bilingual preschool Singaporean children. *International Journal of Language and Communication Disorders, 49*(3), 317-332. doi:10.1111/1460-6984.12075

Gopnik, M., & Crago, M. B. (1991). Familial aggregation of a developmental language disorder. *Cognition, 39*(1), 1-50. doi: 10.1016/0010-0277(91)90058-C

Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*(1), 177-195. doi:10.1044/1092-4388(2007/015)

Gray, S. (2003). Word-learning by preschoolers with specific language impairment: What predicts success? *Journal of Speech, Language, and Hearing Research, 46*(1), 56-67. doi:10.1044/1092-4388(2003/005)

Gray, S. (2004). Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research, 47*(5), 1117.

Gray, S. (2005). Word learning by preschoolers with specific language impairment: Effect of phonological or semantic cues. *Journal of Speech, Language, and Hearing Research, 48*(6), 1452-1467. doi:10.1044/1092-4388(2005/101)

Gray, S., Pittman, A., & Weinhold, J. (2014). Effect of phonotactic probability and neighborhood density on word- learning configuration by preschoolers with typical development and specific language impairment. *Journal of Speech, Language, and Hearing Research, 57*(3), 1011-1025. doi:10.1044/2014_JSLHR-L-12-0282

Guiberson, M., & Atkins, J. (2012). Speech-language pathologists' preparation, practices, and perspectives on serving culturally and linguistically diverse children. *Communication Disorders Quarterly, 33*(3), 169-180. doi:10.1177/1525740110384132

Gupta, A. F. (1994). *The step-tongue: Children's English in Singapore*. Clevedon: Multilingual Matters.

Gupta, A. F., & Chandler Yeo, H. (1994). Speech and language therapy in Singapore. In A. F. Gupta (Ed.), *The step-tongue: Children's English in Singapore* (pp. 191-205). Clevedon: Multilingual Matters.

Gutiérrez-Clellen, V. F., & Peña, E. (2001). Dynamic assessment of diverse children: A tutorial. *Language Speech and Hearing Services in Schools, 32*(4), 212-224. doi:10.1044/0161-1461(2001/019)

Gutiérrez-Clellen, V. F., & Simon-Cereijido, G. F. (2009). Using language sampling in clinical assessments with bilingual children: Challenges and future directions. *Seminars In Speech And Language, 30*(04), 234-245. doi:10.1055/s-0029-1241722

Håkansson, G. (2001). Tense morphology and verb- second in Swedish l1 children, l2 children and children with SLI. *Bilingualism: Language and Cognition, 4*(1), 85-99. doi:10.1017/S1366728901000141

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2013). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching and Therapy, 29*(1), 57-75. doi:10.1177/0265659012459526

Hasson, N., Dodd, B., & Botting, N. (2012). Dynamic assessment of sentence structure (DASS): Design and evaluation of a novel procedure for the assessment of syntax in children with language impairments. *International Journal of Language & Communication Disorders, 47*(3), 285-299. doi:10.1111/j.1460-6984.2011.00108.x

Hasson, N., & Joffe, V. (2007). The case for dynamic assessment in speech and language therapy. *Child Language Teaching and Therapy, 23*(1), 9-25. doi:10.1177/0265659007072142

Haukoos, J. S., & Lewis, R. J. (2005). Advanced statistics: Bootstrapping confidence intervals for statistics with "difficult" distributions. *Academic Emergency Medicine, 12*(4), 360-365. doi:10.1197/j.aem.2004.11.018

Haywood, H. C., & Lidz, C. S. (2006). *Dynamic assessment in practice: Clinical and educational applications*: Cambridge University Press.

Hemsley, G., Holm, A., & Dodd, B. (2014). Identifying language difference versus disorder in bilingual children. *Speech, Language and Hearing, 17*(2), 101-115. doi:10.1179/2050572813Y.0000000027

Ho, L. Y. (2007). Child development programme in Singapore 1988 to 2007. *Annals Academy of Medicine, 36*(11), 898-910.

Hoff, E., & Core, C. (2013). Input and language development in bilingually developing children. *Seminars In Speech And Language, 34*(4), 215. doi:10.1055/s-0033-1353448

Hoff, E., Core, C., Place, S., Rumiche, R., Senor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*(1), 1-27. doi:10.1017/S0305000910000759

Jackson, E., Leitao, S., & Claessen, M. (2016). The relationship between phonological short-term memory, receptive vocabulary, and fast mapping in children with specific language impairment. *International Journal of Language & Communication Disorders, 51*(1), 61-73. doi:10.1111/1460-6984.12185

Jordaan, H., Shaw-Ridley, G., Serfontein, J., Orelowitz, K., & Monaghan, N. (2001). Cognitive and linguistic profiles of specific language impairment and semantic-pragmatic disorder in bilinguals. *Folia Phoniatrica et Logopaedica, 53*(3), 153-165. doi:10.1159/000052670

Kane, M. T. (1992). An argument-based approach to validation. *Psychological Bulletin, 112*, 527-5235.

Kane, M. (2012). Validating score interpretations and uses. *Language Testing, 29*(1), 3-17. doi:doi:10.1177/0265532211417210

Kane, M. T. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (4 ed., pp. 17-64). Westport, CT: American Council on Education and Praeger.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1-73. doi:10.1111/jedm.12000

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech and Hearing Services in Schools, 43*(1), 81-96. doi:10.1044/0161-1461(2011/10-0095)

Kelly, D. J. (2010). Language acquisition challenges for preschoolers residing in low-ses households: Implications for speech-language pathologists and developmental researchers. *Perspectives on Language Learning and Education, 17*(2), 41-48. doi:10.1044/lle17.2.41

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders, 43*(6), 456-473. doi:10.1016/j.jcomdis.2010.02.002

Kohnert, K., & Windsor, J. (2004). The search for common ground. *Journal of Speech, Language, and Hearing Research, 47*(4), 891-903.doi:10.1044/1092-4388(2004/066)

Kohnert, K., Windsor, J., & Ebert, K. D. (2009). Primary or specific language impairment and children learning a second language. *Brain and Language, 109*(2), 101-111. doi:10.1016/j.bandl.2008.01.009

Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-based processing tasks separate children with language impairment from typical bilinguals? *Learning Disabilities, Research and Practice, 21*(1), 19-29. doi:10.1111/j.1540-5826.2006.00204.x

Laing, S. P., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech and Hearing Services in Schools, 34*(1), 44-55.

Law, N. C. W., & So, L. K. H. (2006). The relationship of phonological development and language dominance in bilingual Cantonese-putonghua children. *International Journal of Bilingualism, 10*(4), 405-427. doi:10.1177/13670069060100040201

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling developmental language difficulties from school entry into adulthood: Literacy, mental health, and employment outcomes. *Journal of Speech, Language, and Hearing Research, 52*(6), 1401-1416. doi:10.1044/1092-4388(2009/08-0142)

Lee, L. C. M., & Rickard Liow, S.J. (2009). *Cognitive Linguistic Assessment Profile*. Unpublished measure. National University of Singapore.

Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge, MA: The MIT Press.

Leonard, L. B., Eyer, J. A., Bedore, L. M., & Grela, B. G. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*(4), 741. doi:10.1044/jslhr.4004.741

Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50*(2), 408-428. doi:10.1044/1092-4388(2007/029)

Lian, W. B., Ho, S. K., Choo, S. H., Shah, V. A., Chan, D. K., Yeo, C. L., & Ho, L. Y. (2012). Children with developmental and behavioural concerns in singapore. *Singapore Med J, 53*(7), 439-445.

Lidz, C. S. (1991). *Practioner's guide to dynamic assessment*. New York: Guilford.

Liiva, C. A., & Cleave, P. L. (2005). Roles of initiation and responsiveness in access and participation for children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*(4), 868-883. doi:10.1044/1092-4388(2005/060)

Lim, V. P. C., Liow, S. J. R., Lincoln, M., Chan, Y. H., & Onslow, M. (2008). Determining language dominance in English–mandarin bilinguals: Development of a self-report classification tool for clinical use. *Applied Psycholinguistics, 29*(3), 389-412. doi:10.1017/S0142716408080181

Low, L. L., Wah, W., Ng, M. J., Tan, S. Y., Liu, N., & Lee, K. H. (2016). Housing as a social determinant of health in Singapore and its association with readmission risk and increased utilization of hospital services. *Frontiers in Public Health, 4*, 109. doi:10.3389/fpubh.2016.00109

McKean, C., Letts, C., & Howard, D. (2013). Triggering word learning in children with language impairment: The effect of phonotactic probability and neighbourhood density. *Journal of Child Language, 41*(6), 1224-1248. doi:10.1017/S0305000913000445

Messick. (1995). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *The American Psychologist, 50*(9), 741. doi:10.1002/j.2333-8504.1994.tb01618.x

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3 ed., pp. 13-103). New York, NY: Macmillan Publishing.

Messier, J., & Wood, C. (2015). Facilitating vocabulary acquisition of children with cochlear implants using electronic storybooks. *Journal of Deaf Studies and Deaf Education, 20*(4), 356-373. doi:10.1093/deafed/env031

Miller, C. A., Kail, R., Leonard, L. B., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(2), 416-433. doi:10.1044/1092-4388(2001/034)

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*(1), 40-48. doi:10.1037/0021-843X.110.1.40

Ministry of Social and Family Development, Singapore. (2013). *Rollout of early intervention programme in pre-schools to support children with mild developmental delays.* Retrieved from http://app.msf.gov.sg/Press-Room/Development-Support-Programme-in-pre-schools

Montgomery, J. (2002). Understanding the language difficulties of children with specific language impairments: Does verbal working memory matter? *American Journal of Speech - Language Pathology, 11*(1), 77. doi:10.1044/1058-0360(2002/009)

Nash, M., & Donaldson, M. L. (2005). Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research, 48*(2), 439-458. doi:10.1044/1092-4388(2005/030)

Oetting, J. B., & Horohov, J. E. (1997). Past- tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*(1), 62-74.

Oetting, J. B., Rice, M. L., & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without sli. *Journal of Speech, Language, and Hearing Research, 38*(2), 434-445. doi:10.1044/jshr.3802.434

Okalidou, A., Syrika, A., Beckman, M. E., & Edwards, J. R. (2011). Adapting a receptive vocabulary test for preschool-aged Greek-speaking children. *International Journal of Language and Communication Disorders, 46*(1), 95-107. doi:10.3109/13682821003671486

Olswang, L. B., & Bain, B. (1996). Assessment information for predicting upcoming change in language production. *Journal of Speech and Hearing Research, 39*(414-423). doi:10.1044/jshr.3902.414

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech and Hearing Services in Schools, 36*(3), 172-187. doi:10.1044/0161-1461(2005/019)

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics, 31*(2), 227-252. doi:10.1017/S0142716409990373

Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language and Hearing Research (Online), 59*(1), 171-182. doi:10.1044/2015_JSLHR-L-15-0008

Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research, 59*(1), 171-182. doi:10.1044/2015_JSLHR-L-15-0008

Paradis, J., & Crago, M. (2001). The morphosyntax of specific language impairment in French: An extended optional default account. *Language Acquisition, 9*(4), 269-300.

Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders, 43*(6), 474-497. doi: 10.1016/j.jcomdis.2010.01.002

Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language

impairment. *Applied Psycholinguistics, 29*(04), 689-722. doi: 10.1017/S0142716408080296

Parsons, S., Schoon, I., Rush, R., & Law, J. (2011). Long-term outcomes for children with early language problems: Beating the odds. *Children & Society, 25*(3), 202-214. doi:10.1111/j.1099-0860.2009.00274.x

Patterson, J. L., Rodríguez, B. L., & Dale, P. S. (2013). Response to dynamic language tasks among typically developing Latino preschool children with bilingual experience. *American Journal of Speech-Language Pathology, 22*(1), 103-112. doi:10.1044/1058-0360(2012/11-0129)

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43*(1), 93-120. doi:10.1111/j.1467-1770.1993.tb00174.x

Peña, E. (1993). *Dynamic assessment: A non-biased approach for assessing the language of young children.* (Doctoral Dissertation).

Peña, E. (1996). Dynamic assessment: The model and language applications. In K. Cole, P. S. Dale, & D. Thal (Eds.), *Assessment of communication and language* (pp. 281-307). Baltimore, MD: Brookes.

Peña, E., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech - Language Pathology, 10*(2), 138-154. doi:10.1044/1058-0360(2001/014)

Peña, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. *The Journal of Special Education, 26*(3), 269-280. doi:10.1177/002246699202600304

Peña, E. D. (2000). Measurement of modifiability in children from culturally and linguistically diverse backgrounds. *Communication Disorders Quarterly, 21*(2), 87. doi:10.1177/152574010002100203

Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *Journal of Speech, Language, and Hearing Research, 57*(6), 2208-2220. doi:10.1044/2014_JSLHR-L-13-0151

Peña, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., & Sabel, T. (2006). Dynamic assessment of school- age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech, Language, and Hearing Research, 49*(5), 1037-1057. doi:10.1044/1092-4388(2006/074)

Peña, E. D., Reséndiz, M., & Gillam, R. B. (2007). The role of clinical judgements of modifiability in the diagnosis of language impairment. *International Journal of Speech-Language Pathology, 9*(4), 332-345. doi:10.1080/14417040701413738

Perrie, S., & Core, C. (2006). *Challenges facing monolingual speech-language pathologists in south Florida.* Paper presented at the ASHA Annual Convention, Miami, Florida.

Petersen, D. B., Chanthongthip, H., Ukrainetz, T. A., Spencer, T. D., & Steeve, R. W. (2017). Dynamic assessment of narratives: Efficient, accurate identification of language impairment in bilingual students. *Journal of Speech, Language, and Hearing Research, 60*(4), 983-998. doi:10.1044/2016_JSLHR-L-15-0426

Phoon, H. S., & Maclagan, M. A. (2009). A survey of Malaysian speech-language pathologists: Perception of articulation and phonological assessments. *Asia Pacific Journal of Speech, Language, and Hearing, 12*(4), 315-332. doi:10.1179/136132809805335256

Pieretti, R. A., & Roseberry-McKibbin, C. (2016). Assessment and intervention for English language learners with primary language impairment. *Communication Disorders Quarterly, 37*(2), 117-128. doi:10.1177/1525740114566652

Prelock, P. A., Hutchins, T., & Glascoe, F. P. (2008). Speech-language impairment: How to identify the most common and least diagnosed disability of childhood. *The Medscape Journal of Medicine, 10*(6), 136-136.

Pua, E. P. K., Lee, M. L. C., & Rickard-Liow, S. J. (2017). Screening bilingual preschoolers for language difficulties: Utility of teacher and parent reports. *Journal of Speech, Language, and Hearing Research, 60*(4), 950-968. doi:10.1044/2016_JSLHR-L-16-0122

Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., . . . Wake, M. (2014). Specific language impairment: A convenient label for whom? *International Journal of Language & Communication Disorders, 49*(4), 416-451. doi:10.1111/1460-6984.12102

Rice, M. L. (2004). Growth models of developmental language disorders. In M. L. W. Rice, S. F. (Ed.), *Developmental language disorders: From phenotypes to etiologies* (pp. 207-240). Mahwah, NJ: Erlbaum.

Rice, M. L., & Hoffman, L. (2015). Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2;6 to 21 years of age. *Journal of Speech, Language, and Hearing Research, 58*(2), 345-359. doi:10.1044/2015_JSLHR-L-14-0150

Rice, M. L., Noll, K. R., & Grimm, H. (1997). An extended optional infinitive stage in german-speaking children with specific language impairment. *Language Acquisition, 6*(4), 255-295. doi:10.1207/s15327817la0604_1

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 37*(1), 106-122. doi:10.1044/jshr.3701.106

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research, 39*(6), 1239-1257. doi:10.1044/jshr.3906.1239

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language, and Hearing Research, 38*(4), 850-863. doi:10.1044/jshr.3804.850

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language and Hearing Research, 41*(6), 1412. doi:10.1044/jslhr.4106.1412

Roseberry-McKibbin, C., & O'Hanlon, L. (2005). Nonbiased assessment of English language learners: A tutorial. *Communication Disorders Quarterly, 26*(3), 178-185. doi:10.1177/15257401050260030601

Sabanayagam, C., Shankar, A., Wong, T. Y., Saw, S. M., & Foster, P. J. (2008). Socioeconomic status and overweight/obesity in an adult Chinese population in singapore. *Journal of Epidemiology, 18*(1), 43-44. doi:10.2188/jea.18.43

Semel, E., Wiig, E. H., & Secord, W. (2003). *Clinical evaluation of language fundamentals fourth edition.* San Antonio, TX: PsychCorp.

Semel, E., Wiig, E. H., & Secord, W. (2004). *Clinical evaluation of language fundamentals preschool edition 2.* San Antonio, TX: PsychCorp.

Sheng, L., Lu, Y., & Gollan, T. H. (2014). Assessing language dominance in mandarin-English bilinguals: Convergence and divergence between subjective and objective

measures. *Bilingualism (Cambridge, England), 17*(2), 364-383. doi:10.1017/S1366728913000424

Singapore Statistics Board. (2010). *Census of population 2010 statistical release 1: Demographic characteristics, education, language and religion*. Singapore Department of Statistic, Ministry of Trade & Industry, Singapore.

Singapore Statistics Board. (2012). *Singapore standard classification of type of dwelling*. Retrieved from https://www.singstat.gov.sg/docs/default-source/default-document-library/methodologies_and_standards/standards_and_classifications/sctd.pdf

Singapore Statistics Board. (2014). *Population trends*. Singapore Department of Statistic, Ministry of Trade & Industry, Singapore.

Speech-Language and Audiology Canada. (1997). *Sac position paper: Speech-language pathology and audiology in the multicultural, multilingual context*. Retrieved from http://www.sac-oac.ca/professional-resources/resource-library/sac-position-paper-speech-language-pathology-and-audiology

Speech Pathology Australia. (2016). *Position Paper: Working in a Culturally and Linguistically Diverse Society.* Retrieved from https://www.speechpathologyaustralia.org.au/spaweb/Document_Management/Public/Position_Statements.aspx

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*: Cambridge University Press.

Stokes, S. F., & Wong, A. M. Y. (1996). Validation of a Cantonese version of the preschool language assessment instrument. *Asia Pacific Journal of Speech, Language and Hearing, 1*(2), 75-90. doi:10.1179/136132896805577496

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language and Hearing Research, 44*(6), 1321. doi:10.1044/1092-4388(2001/103)

Storkel, H. L., & Adlof, S. M. (2009). The effect of semantic set size on word learning by preschool children. *Journal of Speech, Language and Hearing Research, 52*(2), 306-320. doi:10.1044/1092-4388(2009/07-0175)

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research, 49*(6), 1175-1192.

Stow, C., & Dodd, B. (2003). Providing an equitable service to bilingual children in the uk: A review. *International Journal of Language and Communication Disorders, 38*, 351-377. doi:10.1080/1368282031000156888

Summers, C., Bohman, T. M., Gillam, R. B., Peña, E., & Bedore, L. M. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language & Communication Disorders, 45*(4), 480-493. doi:10.3109/13682820903198058

Sze, W.P., & Rickard Liow, S.J. (2009). *Bilingual language assessment battery (BLAB)*. Unpublished measure. National University of Singapore.

Tan, S. H. (2008). *Language background questionnaire IV*. National University of Singapore.

Teoh, W. Q., Brebner, C., & McAllister, S. (2017). Bilingual assessment practices: Challenges faced by speech-language pathologists working with a predominantly bilingual population. *Speech, Language and Hearing*, 1-12. doi:10.1080/2050571X.2017.1309788

Teoh, W. Q., Brebner, C., & McCormack, P. (2012). Assessing the language skills of children from culturally and linguistically diverse backgrounds: The expressive vocabulary performance of Singaporean English- mandarin bilingual pre- schoolers. *International Journal of Speech-Language Pathology, 14*(3), 281-291. doi:10.3109/17549507.2011.652171

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism, 15*(4), 426-445. doi:10.1177/1367006911403202

Thordardottir, E. (2015). The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology, 17*(2), 97-114. doi:10.3109/17549507.2014.923509

Thordardottir, E., & Brandeker, M. (2012). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders*. doi:10.1016/j.jcomdis.2012.08.002

Thordardottir, E., Kehayia, E., Mazer, B., Lessard, N., Majnemer, A., Sutton, A., . . . Chilingaryan, G. (2011). Sensitivity and specificity of French language and processing measures for the identification of primary language impairment at age 5. *Journal of Speech, Language, and Hearing Research, 54*(2), 580-597. doi:10.1044/1092-4388(2010/09-0196)

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*(6), 1245-1260. doi:doi:10.1044/jslhr.4006.1245

Ukrainetz, T. A., Harpell, S., Walsh, C., & Coyle, C. (2000). A preliminary investigation of dynamic assessment with Native American kindergartners. *Language, Speech and Hearing Services in Schools, 31*(2), 142-154.

Van Teijlingen, E. R., & Hundley, V. (2002). The importance of pilot studies. *Nursing Standard, 16*(40), 33-36. doi:10.7748/ns2002.06.16.40.33.c3214

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language, 68*(1-2), 306-311. doi:10.1006/brln.1999.2116

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: MIT Press.

Williams, C. J., & McLeod, S. (2012). Speech-language pathologists' assessment and intervention practices with multilingual children. *International Journal of Speech-Language Pathology, 14*(3), 292-305. doi:10.3109/17549507.2011.636071

Williams, K. T. (2007). *Expressive vocabulary test second edition.* San Antonio, TX: Pearson.

Yeo, S. X. D. (2015). *Mean length of utterance in both languages of typically developing bilingual children: Preliminary data from Singapore.* Unpublished Honours Thesis. Flinders University.

Yeoh, B., & Lin, W. Q. (2012). *Rapid growth in Singapore's immigrant population brings policy challenges.* Retrieved from http://www.migrationpolicy.org/article/rapid-growth-singapores-immigrant-population-brings-policy-challenges

Yip, V., & Matthews, S. (2006). Assessing language dominance in bilingual acquisition: A case for mean length utterance differentials. *Language Assessment Quarterly, 3*(2), 97-116. doi:10.1207/s15434311laq0302_2

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool language scales edition 4.* San Antonio, TX: PsychCorp.

Appendices

Appendix 1: Original Article (Bilingual assessment practices: Challenges faced by speech-language pathologists working with a predominantly bilingual population)

Appendix 2: Ethics Approval for Phase One

Appendix 3: Survey Questionnaire: Assessment practices and challenges faced in assessing the oral language skills of Singaporean bilingual children

Appendix 4: DA of word learning skills scoring sheet

Appendix 5: Ethics Approval for Phase Two (SBREC)

Appendix 6: Ethics Approval for Phase Two (Singhealth)

Appendix 7: Participant Information Sheet

Appendix 8: Letter to Parents

Appendix 9: Language Background Questionnaire

Appendix 10: Articulation Screener

Appendix 1: Original Article (Bilingual assessment practices: Challenges faced by speech-language pathologists working with a predominantly bilingual population)

This manuscript has been accepted for publication in Speech, Language and Hearing.

Wei Qin Teoh, Chris Brebner & Sue McAllister (2017): Bilingual assessment practices: challenges faced by speech-language pathologists working with a predominantly bilingual population, *Speech, Language and Hearing*, DOI: 10.1080/2050571X.2017.1309788

Bilingual assessment practices: Challenges faced by speech-language pathologists working with a predominantly bilingual population

Wei Qin Teoh[1,2] , Chris Brebner[1], Sue McAllister[1]

[1]*Speech Pathology and Audiology, Flinders University, Adelaide, Australia*

[2]*Department of Child Development, KK Women's and Children's Hospital, Singapore*

Speech Pathology
Level 7 Flinders Medical Centre
School of Health Sciences
Flinders University
GPO Box 2100
Adelaide SA 5001

Email: teoh0024@flinders.edu.au

Bilingual assessment practices: Challenges faced by speech-language pathologists working with a predominantly bilingual population

The oral language assessment of bilingual children is challenging. The assessment practices and challenges faced by speech-language pathologists (SLPs) in countries where bilingualism is the norm have not been well investigated. This paper summarises what is known about recommended bilingual assessment measures and their limitations. This leads to the investigation of the assessment measures used and challenges faced by SLPs working in Singapore, an English speaking and predominantly bilingual country. SLPs working with children in Singapore were invited to participate in an online survey that centred on the themes of assessment practices and challenges via email invitations. A total of 26 responses were analysed. Results indicated that although the majority of the SLPs were bilingual, they too faced many challenges in assessing bilingual children's language skills. The lack of appropriate local assessment tools, data on the developmental trajectories of local languages and, the lack of practice guidelines on bilingual assessment and alternative measures have resulted in SLPs using standardised assessments that were not designed for use with the population. Despite recommendations from the literature, there was also inadequate use of alternative assessment measures. Given the diversity of bilingual children's language background and development, alternative assessments should be further explored to evaluate their skills rather than further efforts to modify or re-norm current validated standardised assessments.

Keywords: bilingual, language assessment, alternative assessment, standardised assessment, Singapore, assessment challenges

Introduction

Valid and reliable assessment of oral language skills is a crucial step in the provision of speech-language pathology (SLP) services for bilingual[4] children (De Lamo White & Jin, 2011; Hemsley, Holm & Dodd, 2014; Laing & Kamhi, 2003). Appropriate assessments enable Speech Language Pathologists (SLPs) to reliably differentiate true language impairment from language difference in bilingual children. As the number of bilingual speakers has increased globally, so has the number of bilingual children seen by SLPs (Hemsley et al., 2014). This paper briefly summarises bilingual assessment measures and extends these insights via a survey of current needs and assessment practices of SLPs working in Singapore, an English speaking but predominantly bilingual country.

Assessment challenges: The reality

Bilingual children's language skills are challenging to assess as SLPs must be able to successfully differentiate between language impairment, which requires intervention, and language difference, which does not. In a heterogeneous bilingual population where a variety of languages are spoken, language proficiency and dominance among bilingual children within the population and even within a bilingual child, can vary over time as a result of changing language experiences and contexts (Hoff & Core, 2013; Hoff & Core, 2015; Gathercole, 2014; Kohnert, 2010; Thordardottir, 2015). The linguistic abilities of bilingual children in each language range from native to limited proficiency throughout their education years depending

[4] In this paper, the term bilingual also refers to the term multilingual. The term bilingual refers to a person who has been exposed to two or more languages at home, school and in the community; and is able to understand and speak two or more languages in everyday conversations.

on a) quality and quantity of exposure and b) aptitude and motivation (e.g. Gathercole, 2014; Hoff & Core, 2013; Kohnert, 2010).

The variability in language experiences and possibility of cross-linguistic influence when learning two or more languages results in a wide range of language abilities among bilingual children even within the same community. Studies have shown that bilingual children acquire structures (i.e. vocabulary, morphosyntax) for each of their languages at a different rate compared to their monolingual counterparts (Brebner, McCormack & Rickard-Liow, 2016; Thordardottir, 2005). This variability makes it challenging for SLPs to use established clinical markers of language impairments based on the language development trajectories of monolingual children to identify language impairment among bilingual children.

Bilingual assessment also presents challenges for SLPs who do not speak the dominant language of the bilingual child. SLPs often speak a limited number of languages compared to the diverse range of languages spoken in bilingual communities they serve. There is often a mismatch with the languages the SLPs and the bilingual children predominantly speak (Caesar & Kohler, 2007; Cruz-Ferreira & Ng, 2010; D'Souza, Bird & Deacon, 2012; Williams & McLeod, 2012). Increasing the number of bilingual SLPs to meet the increasing demand to provide SLP services in linguistically diverse communities is frequently proposed as a solution (D'Souza et al., 2012; Perrie & Core, 2006). However, even SLPs who can speak the child's languages need culturally and linguistically appropriate assessment tools (Carter, Lees, Murira, Gona, Neville & Newton, 2005). Culturally and linguistically appropriate assessments and information on the developmental trajectories of local languages spoken and/or information to discriminate between language impairment and difference in bilingual communities are rarely available (Brebner, Chandler Yeo, Goh, Kam & Yeo, 2015; Brebner et al., 2016; Guiberson & Atkins; 2012; Stow & Dodd, 2003; Teoh, Brebner & McCormack, 2012; Williams & McLeod, 2012).

A range of assessment measures used by SLPs working with children from bilingual communities in predominantly English-speaking monolingual countries is briefly summarised in the next section. These assessment measures may not yield a valid judgement of the bilingual child's language ability if they are not informed by other critical information. This includes the child's performance in all the languages he/she speaks while taking into consideration information about his/her language environment, patterns of language use and language dominance, first language abilities and developmental history gathered through case history interviewing, parent report and/or standardised questionnaires (Brebner et al, 2016; Bedore et al., 2012; De Lamo White & Jin, 2011; Paradis, Emmerzael & Duncan, 2010; Paradis, 2016; Thordardottir, 2015).

Standardised assessments

International surveys of SLPs' assessment practices with bilingual children (e.g. Caesar & Kohler, 2007; D'Souza et al., 2012; Williams & McLeod, 2012) found that standardised assessments are a common choice among SLPs when assessing bilingual children. Standardised assessments allow SLPs to evaluate a child's existing knowledge in one language and compare it to same age peers who are typically developing. Most standardised assessments are created and standardised for specific populations, usually monolingual speakers of the language. These assessments can only yield valid judgements if the bilingual child has similar language experience and exposure to the normative sample (De Lamo White & Jin, 2011). However, this is often not the case as language experience and/or exposure among bilingual children are often variable. The contents of these standardised assessments also do not match the cultural and linguistic experiences of bilingual children. Therefore, the resulting scores cannot be meaningfully interpreted or used to judge between typical developing or an impaired performance (Kohnert, 2010; Teoh et al., 2012).

Assessment adaptation/modification and re-norming have been explored to extend the use of popular standardised assessments with bilingual populations, (e.g. Brebner, 2010; Okalidou, Syrika, Beckman & Edwards, 2011; Stokes & Wong, 1996). Although content and pictorial modifications may make standardised assessments more culturally and linguistically appropriate, the collection of normative data for the bilingual population is not only time-consuming but also impractical. It also requires a relatively homogenous group (De Lamo White & Jin, 2011) whereas bilingual populations are more often heterogeneous in nature. The variability in language experience and dominance within each bilingual population leads to different rates and patterns of language acquisition and development even among same age children from the same ethnic group (Brebner et al., 2015; Brebner et al, 2016; Teoh et al, 2016). Even if normative data are collected and available, it is impossible to determine if the bilingual child has the same language experience and exposure as the normative sample to be able to make a meaningful comparison.

Conceptual scoring (also known as composite scoring) has been employed as an alternative to single language scoring when assessing bilingual children on standardised assessments (Pearson, Fernandez & Oller, 1993; Gross, Buac & Kaushanskaya, 2014). Conceptual scoring counts the total number of vocabulary and basic concepts a bilingual child can express in any of the languages he or she is exposed to (Gross et al., 2014). Conceptual scoring is found to improve the scores of bilingual children to be comparable to that of monolingual children (Pearson et al., 1993; Thordardottir, Rothenberg, Rivard & Naves, 2006). It also reduces the over-identification of language impairment among simultaneous bilingual children (Gross et al., 2014). However, meaningful interpretation of conceptual scores is still problematic. Standardised assessments are designed to assess a single language hence translating test items into other languages for conceptual scoring purposes will not preserve the psychometric structure of the assessment (e.g. item difficulty hierarchy and the basal and ceiling rules of the

standardised assessment) (Gross et al., 2016). Secondly, standardised assessment designed for specific populations can have culturally and linguistically bias items (De Lamo White & Jin, 2011). For example, in a standardised assessment designed for use in the USA, the picture of a 'totem pole' is familiar to American children but is not familiar to children from non-American cultures. In addition, translated equivalents in other languages do not also always exist.

De Lamo White and Jin (2011) have recommended that when standardised assessments are used, they should not be scored but instead used to provide a description of a bilingual child's language ability. However, this alone still does not provide sufficient information to allow SLPs to judge whether the bilingual child does or does not have a language impairment. Paradis and colleagues (2010, 2016) recommend also using parent reports and/or standardised questionnaires to provide further information on the bilingual children's language development and ability to assist in accurate identification of language impairment.

Criterion-referenced measures

A criterion-referenced measure is an alternative assessment approach that evaluates a bilingual child's existing knowledge in a specific language domain (i.e. grammatical structure, linguistic concept) in comparison with a predetermined performance criterion. The advantage of criterion-referenced measures is that SLPs can use materials in interaction patterns and contexts that are familiar to the bilingual children hence minimising cultural and linguistic bias (De Lamo White & Jin, 2011; Laing & Kamhi, 2003).

The predetermined criterion is based on a set of language performances that is expected of bilingual children from a specific community/population. This predetermined criterion is usually derived from known language development/acquisition data for that language spoken in the bilingual community (De Lamo White & Jin, 2011; Laing & Kamhi, 2003). However, in

reality, there is often a lack of well-established language developmental data for bilingual populations/communities. As such, it is often difficult for SLPs to set a valid predetermined criterion to determine whether the bilingual child is performing like his/her peers in a specific language domain or has a core language difficulty.

To overcome the lack of well-established developmental data, some researchers have suggested that the predetermined criterion to be determined based on language patterns of parents or caregivers (Parent-Child Comparative Analysis: Terrell, Arensberg & Rosa, 1992). However, it has been argued that it is highly inaccurate to assume that linguistic behaviours of a child can match to an adult (Laing & Kamhi, 2003). Besides the study conducted by Terrell and colleagues (1992), there is no other bilingual study to date that investigated the reliability of using predetermined criterions that are based on the language patterns of parents or caregivers.

The research on using criterion-referenced measures to identify language impairment in bilingual children is still sparse. It is recommended that criterion-referenced measures should not be used alone. Instead, they are to be supplemented with background information (i.e. parent report and/or parent questionnaire) and the use of additional/other assessment measures before determining if the bilingual child has or does not have a language impairment (De Lamo White & Jin, 2011; Laing & Kamhi, 2003).

Language sampling

Language sampling is one of the more common alternative assessment approaches adopted by SLPs assessing bilingual children (D'Souza et al, 2012; Williams & McLeod, 2012). Language samples can be collected through narratives, picture description, and spontaneous conversations while using language, materials and interaction patterns that are familiar to the bilingual child (De Lamo White & Jin, 2011; Laing & Kamhi, 2003). Language samples can provide information on the child's language functioning, lexical knowledge and use of

linguistic structures in both languages, for comparison to data on language development for the languages of the bilingual community (e.g. Gutiérrez-Clellen & Simon-Cereijido, 2009; Ooi & Wong, 2012). For example, Ooi and Wong (2012) found that utterance length and syntax production information obtained from language sampling has the potential to provide diagnostic information to identify language impairment in Chinese-English bilingual Malaysian children.

It is recommended that language samples should be collected in both the languages a bilingual child speaks (Gutiérrez-Clellen & Simon-Cereijido, 2009). SLPs can refer to established clinical behavioural markers of language impairment or established language acquisition/development data for each language to assist in decision making. Bilingual children with language impairment should demonstrate difficulties in not one but both languages they speak.

There are, however, challenges in using language sampling as an assessment measure. In reality, besides the English language, there is still a lack of research on the clinical behavioural markers of language impairment and developmental/acquisition data for other languages, especially for non-English languages for SLPs to make a comparison to (Bedore & Peña, 2008; Gutiérrez-Clellen & Simon-Cereijido, 2009; Ooi & Wong, 2012). Even if monolingual developmental data for the language is available, some bilingual children, especially those with limited exposure to a language, may demonstrate clinical behavioural markers of language impairment in that language that is related to language differences as a result of cultural, linguistic and environmental influences (Paradis, 2005; Thordardottir, 2015). In the absence of an interpreter, language sampling analysis can also be challenging for SLPs who do not speak either or both languages spoken by the bilingual child.

Processing-dependent measures

Processing-dependent measures are not a common alternative assessment measure used by SLPs despite being proposed as a least biased measurement (Williams & McLeod, 2012). Language processing measures evaluate children's underlying cognitive processing abilities that are used for processing and learning languages as well as other cognitive operations. Unlike the other assessment measures (e.g. standardised assessments, criterion-referenced measures, language sampling) which evaluate the children's existing language knowledge, processing functions that are not influenced by existing language knowledge are being evaluated instead. Some examples of language processing measures are working memory tasks (e.g. digit recall, non-word repetition, word learning), perceptual tasks (e.g. discrimination of rapidly presented tones) and competing stimuli tasks (e.g. auditory figure ground, competing language processing task) (De Lamo White & Jin, 2011; Kohnert, Windsor & Yin, 2006; Laing & Kamhi, 2003).

Language processing measures hold promise as a less biased assessment approach to identify language impairment in bilingual children. However, findings (De Lamo & White, 2011; Kohnert et al., 2006) have been mixed as to whether poor performances in language processing tasks such as sentence repetition, non-word repetition can validly be interpreted as associated with, or predictive of, language impairment in bilingual children. There is, however, some emerging evidence that performance in word learning tasks can discriminate between bilingual children with language impairment and typically developing bilingual children (Kapantzolgou, Restrepo & Thomspon, 2012). Nonetheless, researchers in the field agree that the diagnostic value of language processing measures to identify language impairment in bilingual children has not been fully explored and more research in this area is required (De Lamo White & Jin, 2011; Laing & Kamhi, 2003; Roseberry-McKibbin & O'Hanlon, 2005).

Dynamic Assessment

Dynamic assessment (DA) is becoming a preferred alternative assessment measure for assessing bilingual children (D'Souza et al, 2012; Willams & McLeod, 2012). DA evaluates the children's capacity to learn language targets in response to learning opportunities provided within the assessment itself. It allows SLPs to observe how children learn the language. It is also known be a less biased assessment approach as unlike other knowledge-based assessment measures (e.g. standardised assessments, criterion-referenced measures, language sampling), children's performances on DA are not affected by the environmental (i.e. low social economic status, lack of schooling) and linguistic variables (i.e. lack of exposure to the language, other varieties of the language).

DA draws on Vygotsky's model of cognitive development which proposes that learning takes place in the "Zone of Proximal Development" (ZPD) (Vygotsky, 1978; cited in Gutiérrez-Clellen & Peña, 2001: 212). For the purposes of language assessment, this is defined as a language target that is developmentally achievable for the child. The goal of DA is to identify if the child is able to demonstrate the ability to learn language targets that are in their ZPD.

There are several DA approaches (for a detailed discussion, see Gutiérrez-Clellen & Peña, 2001). To discriminate language difference from language impairment, the 'test-teach-retest' paradigm is often used. In the 'test-teach-retest' paradigm, the SLP identifies a language area with which the child has difficulties and provides intervention in the form of mediated learning experiences or graduated prompting to improve the child's functioning in the targeted area. The SLP then evaluates the child's performance again on the targeted area to identify change that would indicate the child's ability to acquire language targets after teaching (Gutiérrez-Clellen & Peña, 2001). Children who show significant improvements are likely to have language difference due to the lack of familiarity or experience with the content assessed,

whereas children who demonstrate limited changes are likely at risk of language impairment.

DA has been shown to be useful in discriminating between typically developing bilingual children from bilingual children with language impairment in vocabulary and word learning studies (Camilleri & Botting, 2013; Kapantzoglou et al., 2012; Peña, Iglesias & Lidz, 2001), narrative studies (Peña et al., 2006) and sentence structure studies (Hasson, Dodd & Botting, 2012).

An advantage of DA is that the assessment can be carried out in a common language (e.g. English) spoken by both the child and SLP. This is especially useful when the SLP does not speak the dominant language of the child. Although developmental data/information for the languages spoken in the bilingual community is not required to make a comparison to, it is useful for SLPs to have knowledge of the developmental trajectories of local languages. This is so that they can select language tasks that can successfully discriminate language impairment from language difference among the bilingual children in the community.

Aims

The assessment measures summarised above are derived from studies and surveys of SLPs' practices with bilingual children within predominantly monolingual countries (Caesar & Kohler, 2007; De Lamo & White, 2011; D'Souza et al, 2012; Stow & Dodd, 2003; Williams & McLeod, 2012). Assessment practices in countries where bilingualism is the norm have not been investigated, and the assessment challenges faced by SLPs in such heterogeneous bilingual communities are also not well understood. This research sought to investigate the nature of and challenges experienced by SLPs assessing children in a country where bilingualism is the norm through a survey. Specifically, we aimed to explore:

Whether SLPs working with a predominantly bilingual population are bilinguals themselves; and if they faced equal or fewer challenges in assessing bilingual children from their community?

SLPs' assessment practices in assessing bilingual children from a predominantly bilingual community.

Challenges faced by SLPs in assessing bilingual children from a predominantly bilingual community.

Methods

Ethical clearance

Ethical clearance for this study was obtained from the Social and Behavioural Research Ethics Committee, Flinders University of South Australia.

Participants' community

The rich language landscape in Singapore and its predominantly bilingual population makes it an appropriate site to investigate the assessment practices and challenges faced by SLPs working in predominantly bilingual populations. Singapore is a multicultural and multilingual country in South-East-Asia. At least 71% of the population (aged 15 and above) are literate in two or more languages (Singapore Statistics Board, 2010). Four official languages (English, Mandarin, Malay, and Tamil) and a variety of other non-official languages are spoken. Bilingual education is compulsory and all content subjects are taught in English, with students studying a second official language based on their ethnicity. Language exposure within the population is highly variable as children can be simultaneously (since birth) or sequentially exposed to two or more languages (first learning one language, followed by another).

Survey Instrument

An online survey was designed based on themes identified in previous surveys on bilingual assessment and intervention practices (Caesar & Kohler, 2007; D'Souza et al., 2012; Guiberson & Atkins, 2012; Williams & McLeod, 2012). The survey comprised questions about demographic information of SLPs (e.g. language proficiency), caseload demographics, current assessment practices and challenges faced.

The survey was piloted with five practicing SLPs working in Singapore who varied in years of practice in Singapore, practice setting, and ethnicity. Feedback resulted in one minor content modification (i.e. creation of two separate questions on assessment practices with English-dominant (ED) bilingual children and non-English dominant (NED) bilingual children) and two minor presentation modifications (i.e. bolding of keywords in questions and reducing the number of questions on a single page). The final questionnaire comprised of 16 questions with Likert scale and closed choice responses. All questions had the option of free text boxes where SLPs could also input their opinions and/or comments.

Accessibility

The survey was made available online through a secure web domain for a nine-week period. Invitations to participate in the survey were sent via e-mail to members of Speech and Language Therapy Singapore (SALTS) through their in-house email bulletin and to the point of contact of all known paediatric SLPs practices in Singapore (hospitals, community services, and private practices – a total of 49 known practices).

Data Analysis

The responses were analysed descriptively using Statistical Program for the Social Sciences (SPSS) for Windows version 22.0.

Results

Participants' demographics

Twenty-seven SLPs working with the paediatric population in Singapore participated anonymously in the survey. Participant #27 was not included in the final analysis as he/she reported that his/her caseload was mostly non-Singaporean children who were monolingual. Thus, only the responses from the remaining 26 SLPs were analysed.

All 26 SLPs reported to be either working in the public healthcare sector or Voluntary Welfare Organizations. No figures for the numbers of SLPs working in paediatric SLP services in these two sectors were available from the Allied Health Professional Council (AHPC) or SALTS. The response rate was calculated by dividing the number of SLPs who participated in the survey with the number of SLPs working in public health care sector and VWOs in Singapore (i.e. total number of SLPs = 237; AHPC Singapore, estimated percentage of SLPs working in paediatric services in Singapore = 49.0%; Cruz-Ferreira and Ng, 2010). The participation rate of SLPs with paediatric caseloads in these two sectors was estimated to be approximately 22.0% *(26/(0.49x237) x 100%)*.

In Table 1, the demographic data (i.e. citizenship, country of SLP qualification) of the 26 SLPs matched to the demographic data of SLPs in Singapore provided in 2014 by the AHPC. All SLPs who participated in the survey reported that the majority of their caseload (Range: 70.0 – 100.0%) were made up of children aged 12 years and below. All SLPs also reported that the children on their caseloads were dominant in a diverse range of languages (Table 2).

Insert Table 1 approximately here

Language background of SLPs and their caseloads' language demographics

Twenty-two (84.6%) of the 26 SLPS spoke at least two languages proficiently[5] or functionally[6]. Of the four official languages spoken in Singapore, only English was spoken proficiently by all, followed by Mandarin (*n*=18; 69.2%). Despite being official languages spoken in Singapore, Malay and Tamil only had one SLP each who spoke the language proficiently or functionally.

Regardless of the language backgrounds of SLPs, the majority had a range of children on their caseloads who were dominant in the different languages spoken in Singapore (Table 2). The proportion of children dominant in the respective languages on the SLPs' caseloads corresponded to the census data on the language demographics of Singaporean children (Singapore Statistics Board, 2010).

Insert Table 2 approximately here

Assessment measures used

SLPs were asked to differentiate the assessment measures used with English-dominant (ED) and non-English dominant (NED) children to identify whether the variability in language experience among bilingual children impacts on the choice of assessment method. In general, standardised assessments were used more frequently than alternative assessments for both groups of children (Table 3). Standardised assessments that were not normed and created for used for the Singaporean population were used more frequently than available local standardised assessment (Table 4). Participants explained their limited use of available local standardised assessment with the following comments: *"not aware of the availability of these local standardised assessments..", "..do not provide sufficient information (i.e. only assess*

[5] Proficiently: Able to use the language fluently and accurately in all social and formal contexts

[6] Functionally: Able to use the language with sufficient fluency and accuracy in most formal and social contexts

vocabulary skills)", "lack of normative data for other ethnic groups" and *"…time-consuming to use"*.

Insert Table 3 approximately here

Insert Table 4 approximately here

For standardised assessments that were not created and normed for use with the Singaporean population, the majority of the SLPs (*n*=23, 84.4%) reported that they would compare the children's performances to normative data in the respective manuals (Table 5). For example, participant #14 justified his/her choice by stating that "*for primary school-age children, their language skills in English should have caught up sufficiently to allow for more accurate comparison against monolingual norms*".

Other methods such as estimating the Singaporean child's performance based on clinical experience or estimating the Singaporean child's performance descriptively by using a task analysis checklist approach were also common (>50.0%).

Insert Table 5 approximately here

Challenges faced by SLPs in providing assessment services to a bilingual population

A list of bilingual assessment challenges known in the literature (D'Souza et al, 2012; De Lamo White & Jin, 2011) combined with assessment challenges faced by the authors working in Singapore was put together for the purpose of the survey. SLPs were asked to rate the frequency with which they encounter each assessment challenge when assessing Singaporean bilingual children. Table 6 shows that eight out of the 14 listed challenges were faced frequently (*all the time*, *often*) by the majority of the SLPs (>*n*=21,>80.8%) in their everyday

clinical practice were associated with the lack of appropriate assessment tools and information of the developmental trajectories of the languages spoken in Singapore.

Insert Table 6 approximately here

Discussion

The purpose of the study was to explore the assessment practices and challenges faced by SLPs working in a predominantly bilingual country. Of particular interest was whether bilingual SLPs faced similar challenges in bilingual assessment practice as their international monolingual counterparts. The other main purposes of the study were to determine whether recommended bilingual assessment practices were being adopted by SLPs working in a predominantly bilingual country and the nature of the assessment challenges faced.

SLPs and their caseload language match

The SLPs in this study were unique because in contrast to past surveys (i.e. Caesar & Kohler: 6.2% of SLPs were bilinguals; Williams & McLeod, 2012: 48.4% of SLPs were bilinguals), the majority of the SLPs who participated (n=22; 84.6%) in this survey spoke at least two or more languages proficiently or functionally. However, despite being bilingual, they still face the same difficulties as their monolingual counterparts in assessing bilingual children as clinician and client languages did not always match.

The SLPs' caseloads were made up of a diverse range of Singaporean bilingual children who spoke different languages. All SLPs reported that they provided assessment and intervention services to a diverse range of children who were dominant in the different languages. For example, participant #2 who spoke English and Mandarin proficiently/functionally reported that 20.0% of his/her caseload comprised of children who were dominant in other languages (i.e. Malay, Tamil). This reflects the heterogeneous nature of bilingual communities. Most of

the SLPs had to assess bilingual children whose dominant language did not match to any of those that they were proficient or functional in.

The results of the study showed that the solution of increasing the number of bilingual SLPs to provide SLP services in linguistically diverse communities alone is not a practical solution. Despite being bilinguals themselves, up to 50.0% (*n*=13) of the SLPs in this study's survey indicated the ability to speak and assess the dominant language of a child on their caseload as an assessment challenge (see Table 6). This demonstrates the fact that although having a common language between clinician and client allows the clinician to assess client's language with less biased, the challenges in assessing bilingual children's language skills still remain.

Assessment practices and challenges

Current assessment practice and the challenges

Even in cases where the languages spoken by the SLP and the bilingual child were matched, SLPs still reported challenges in assessing bilingual children. These challenges stem from the lack of 1) information of the developmental trajectories of local languages (Table 6), 2) appropriate local assessment tools (Table 6), and 3) local guidelines on recommended bilingual assessment practices and alternative measures.

The lack of available information on the developmental trajectories of local languages and appropriate local assessments tools had resulted in SLPs selecting standardised assessments that were not normed and created for use with Singaporean bilingual children. Regardless of the Singaporean bilingual child's language dominance, the use of such standardised assessments was still a preferred choice (Table 3 and 4). This result was similar to that of a survey study by Caesar and Kohler (2007) where they found that SLPs in the USA continued to use standardised assessments that were designed and normed for use with monolingual English-speaking children over alternative assessments when assessing bilingual children.

Even though the linguistic and cultural diversities in Singapore and USA are different, the prevalent use of standardised assessments to assess the language skills of bilingual children are of similar practice. This shows that the assessment challenges remained the same and are a world-wide problem.

From the survey results, SLPs in Singapore reported that they faced multiple challenges when standardised assessments that are designed for monolingual English-speaking populations (i.e. the UK or the USA) were used. This was not a surprise as previous papers (Bedore & Peña, 2008; De Lamo White and Jin, 2011; Kohnert, 2010) had repeatedly highlighted the limitations and challenges that would arise when assessments tools that were designed for use for monolingual speakers of the language are used to assess the language skills of bilingual children. SLPs in Singapore reported the lack of appropriate local normative data (*n*=24, 92.3%), the presence of linguistic bias in standardised assessments (*n*=23, 88.5%) and cultural bias in standardised assessments (*n*=24, 92.3%) as top assessment challenges (Table 6).

The challenges reported by the SLPs indicated that they were aware of the limitations of using standardised assessments that were not designed for the Singaporean population. However, because there are no local normative data, many of them continued to interpret Singaporean children's performance by comparing their performance to the UK or the USA normative data that are available in the manuals (Table 5). Many also reported that they would also attempt to use their clinical judgement (*n*=21, 80.8%) and/or task analysis checklists (*n*=13, 50.0%) to interpret Singaporean bilingual children's performances on standardised assessments. Nonetheless, such methods do not allow SLPs to appropriately determine whether the Singaporean child assessed is typically developing, has a language impairment or a language difference. This demonstrated that although the SLPs in Singapore were challenged by the many assessment issues on the ground, they were still trying their best to manage to ensure quality speech and language services were still being provided.

Why recommended alternative assessments were not as frequently used?

Many researchers (e.g. De Lamo & White, 2011; Hemsley et al., 2014; Laing & Kamhi, 2003) recommend the use of alternative assessments (i.e. DA, language sampling) over standardised assessments when assessing bilingual children. Recent surveys by Williams and McLeod (2012) in Australia and D'Souza and colleagues (2012) in Canada found that the SLPs were using alternative assessments more frequently than standardised assessments when assessing bilingual children. However, this was not the case in Singapore (Table 3). This was likely due to lack of local guidelines on bilingual assessment practices and alternative assessments.

Internationally, position papers and guidelines on bilingual assessment practices (American Speech-Language Hearing Association, 2013; Royal College of Speech and Language Therapists, 2015; Speech-Language and Audiology Canada, 1997; Speech Pathology Australia, 2016) have been published to create awareness and understanding among SLPs in providing equitable speech language services to bilingual children. The increasing awareness in recommended bilingual assessment practices with emphasis in using alternative assessments are starting to be translated into actual clinical practices. For example, recent survey studies (Williams and McLeod, 2012; D'Souza et al, 2012) found that the SLPs were choosing alternative assessments over standardised assessments when assessing bilingual children as compared to 10 years ago (i.e. Caesar & Kohler, 2007).

Besides the lack of practice guidelines in Singapore, there is no local study that has investigated the use of alternative measures to differentiate between typically developing Singaporean bilingual children and Singaporean bilingual children with language impairment. As such, SLPs in Singapore were probably not confident in adopting alternative measures in assessing Singaporean bilingual children and tended to over-rely on standardised assessments as they are known to be "tools of the profession" – (Stow & Dodd, 2003). In addition, the lack of information of the developmental trajectories of local languages and characteristics of language

impairment in Singaporean bilingual children may have contributed to the infrequent use of alternative assessments. This is because some alternative assessments do require SLPs to have information of the developmental trajectories for the languages spoken in the bilingual population. For example, to select a pre-determined criterion (i.e. criterion-referenced measures) or compare a bilingual child's performance to established developmental norms (i.e. language sampling).

Clinical implications

Bilingual children now represent an increasing proportion of speech-language therapy caseloads world-wide (Hemsley et al., 2014). Increasing the number of bilingual SLPs has been suggested to meet the increasing demand to provide SLP services in linguistically diverse (bilingual) communities (e.g. D'Souza et al., 2012; Perrie & Core, 2006). However, challenges in assessing bilingual children, such as mismatch in SLP and client language, will still remain. Results from this study demonstrated that bilingual SLPs too faced similar challenges to their monolingual international counterparts in assessing bilingual children from culturally and linguistically diverse communities. The solution of increasing the number of bilingual SLPs globally alone will not resolve the challenges faced in assessing bilingual children.

The lack of appropriate local assessment tools, information of the developmental trajectories of local languages and, local guidelines on recommended bilingual assessment practices and alternative measures resulted in SLPs using standardised assessments that were not designed and normed for use with the bilingual population. Although solutions such as modifying the assessment to be culturally and linguistically appropriate, 're-norming' and conceptual/composite scoring have been suggested, they may not be most appropriate and effective (De Lamo White & Jin, 2011). This is because bilingual populations throughout the world, even within a bilingual community, are often heterogeneous in nature. It is almost

impossible to gather normative data for a heterogeneous bilingual population. Bilingual children's language abilities also change over time as a result of shifting language experiences and individual differences. Brebner and colleagues (2015) have also highlighted that developing local normative data is often complex for a bilingual population due to the variability of language dominance and language experiences. Furthermore, they also identified that patterns of language use in many bilingual communities are changing rapidly over time, impacting on the validity of such data. As Gathercole (2014) aptly described the difficulty in gathering normative data on bilingual children that "variation is the norm, not the exception" (p.360) and "one size does not fit all" (p.359).

To contribute to the existing knowledge in the literature regarding recommended bilingual assessment practices, the authors of this paper would like to emphasised the following points:

1. When assessing bilingual children, SLPs must be aware of the limitations standardised assessments that were designed for use with monolingual speakers of the language. If such standardised assessments have to be used, it should be used as part of an assessment battery that also includes alternative measures. SLPs must also collect information on language exposure, language ability, and developmental history through parent report/questionnaire as it is known to increase accurate identification of language impairment among bilingual children.
2. Guidelines or/and position papers on recommended bilingual assessment practices and intervention in countries (i.e. Singapore) when bilingualism is the norm should be developed. This may reduce the over-adoption of monolingual assessment practices and create increased awareness of recommended bilingual assessment practices. In the meantime, SLPs can refer to literature on recommended bilingual assessment practices and alternative measures (i.e. Bedore & Peña, 2008; De Lamo White & Jin, 2011; Kohler, 2010; Paradis, 2016; Pieretti & Roseberry-McKibbin, 2016).

3. There is a need to increase the awareness and the use of alternative assessments to assess the language skills of bilingual children among SLPs who work in a predominantly bilingual country/community (i.e. Singapore). As part of evidence-based practices, exploratory studies on the use of alternative assessments to differentiate between typically developing bilingual children and bilingual children with language impairment can be conducted in such populations. With the information and evidences from such studies, SLPs working in predominantly bilingual countries can then be confident in adopting alternative assessments into their clinical practices.
4. Longitudinal studies of the developmental trajectories of local languages spoken in countries (i.e. Singapore) where bilingualism is the norm need to be conducted. These may provide information on possible clinical behavioural markers for language impairments among bilingual children within the community.

Limitations and future research

The results obtained from this study provide some insights into the assessment practices and challenges faced by bilingual SLPs working with a predominantly bilingual population, in this case, in Singapore. However, due to the small estimated response rate and small sample size (22.0%, n=26), and the exploration of the experiences in only one predominantly bilingual country, the results cannot be generalised to the assessment practices of all SLPs in Singapore or globally. Moreover, respondents to the survey only came from two sectors of services in Singapore. According to AHPC data (2014), up to 27.0% of registered SLPs in Singapore are working in the private sector. Although they were approached, no SLP from the private sector participated in the survey study.

SLPs in this study did not list the reasons why they prefer certain assessment measures over others. As such there could be other factors which were not explored (e.g. models of service delivery in Singapore) that could have affected SLPs' choice of assessments. In addition, future research could investigate the assessment practices of SLPs working in other predominantly bilingual countries and compare the assessment practices and challenges faced.

Conclusion

Overall, the results of this study suggest that bilingual SLPs face similar challenges to their monolingual international counterparts when assessing bilingual children. SLP and client languages often did not match, even when the SLP is bilingual. The solution of increasing bilingual SLPs to meet the increasing demand for bilingual speech-pathology services alone is not practical as even bilingual SLPs will face challenges in assessing bilingual children.

The over-reliance on commercially-available standardised assessments to assess bilingual children's language, despite being widely discouraged in literature, can be attributed to 1) lack of appropriate local assessment measures, 2) little or no information on the development trajectories of local languages and, 3) lack of local guidelines on recommended bilingual assessment practices and alternative measures. Recent studies (e.g. Hasson et al., 2013, Hemsley et al., 2014; Kapantzoglou et al, 2012) found that alternative assessments such as DA can accurately language impairment from language difference among bilingual children compared to commercially-available standardised assessments. Globally, there is an ongoing need to translate such research outcomes into actual clinical practice, especially in countries/communities where bilingualism is the norm. To do so and as part of evidence-based practices, it will be worthwhile to conduct studies in countries/communities where bilingualism is the norm to explore the use of alternative assessment measures to discriminate between typically developing bilingual children and bilingual children with language impairment.

Acknowledgement

The authors would like to thank Speech and Language Therapy Singapore (SALTS) for assistance in publicising the survey to its members, and the SLPs who took time from their busy caseloads to participate.

Table 1: Summary of participants' demographic characteristics compared to AHPC's data

Characteristics	Current study *n (%)*	AHPC, 2014 *n* (%)
Citizenship		
Singaporean/Singaporean PR	20 (77.0)	259 (72.0)
Expatriate	5 (19.2)	99 (28.0)
Not stated	1 (3.8)	-
Years of experience as a SLP in Singapore		
<1 year	4 (15.4)	NDA
1-3 years	9 (34.6)	NDA
4-6 years	8 (30.8)	NDA
7-10 years	4 (15.4)	NDA
11-15 years	0 (0.0)	NDA
>15 years	0 (0.0)	NDA
Not stated	1 (3.8)	NDA
Sector of employment		
Public Healthcare Sector (restructured hospitals)	21 (80.8)	173 (48.3)
Government-run Voluntary Welfare Organisations (VWOs)	5 (19.2)	64 (17.9)
Community hospitals	0 (0.0)	19 (5.3)
Private schools, clinics, hospitals, centres	0 (0.0)	96 (26.8)
Educational institutes	0 (0.0)	1 (0.3)
Others	0 (0.0)	5 (1.4)
Sector of previous employment (if any)*		
Public Healthcare Sector (restructured hospitals)	19 (73.1)	NDA
Private hospital/clinic/school	5 (19.2)	NDA
Government-run Voluntary Welfare Organizations (VWOs)	10 (38.5)	NDA
Others	0 (3.8)	NDA

Country of SLP qualification		
Australia	12 (46.2)	153 (42.7)
India	1 (3.8)	39 (10.9)
Ireland	4 (15.4)	23 (6.4)
Singapore	5 (19.2)	55 (15.4)
UK	1 (3.85)	35 (9.8)
USA	0 (0.0)	14 (3.9)
Others (New Zealand, Canada, not stated etc)	3 (11.5)	39 (10.9)

*Some participants have worked in at least one or more other settings previously

NDA: No Data Available

Table 2: Language profiles of SLP participants and language demographics of Singaporean children on their caseload

Language	SLPs' proficiency (proficient / functional) in: *n* (%)	SLPs who see children on their caseload that were dominant in: *n* (%)	Percentage average of children (Age 12 and below) on SLPs' caseloads who were dominant in:	Percentage (Singapore Department of Statistic, 2010) of Singaporean children (ages 5-9) who are dominant in:
English*	26 (100.0)	26 (100.00)	54.3	50.5
Mandarin*	18 (69.2)	25 (96.2)	29.2	28.3
Malay*	1 (3.8)	25 (96.2)	11.9	13.1
Tamil*	1 (3.8)	19 (73.0)	3.9	4.1
Others	3 (11.4)	6 (23.1)	<1.0	NDA

*official languages spoken in Singapore

NDA: No Data Available

Table 3: Comparison of the types of assessments used by SLPs when assessing English dominant (ED) and non-English dominant (NED) Singaporean bilingual children

**Language samples in both languages*

	Assessing the language skills of Singaporean children who are ED			Assessing the language skills of Singaporean children who are NED		
Frequency	All the time/Often *n* (%)	Sometimes *n* (%)	Rarely/ Not at all *n* (%)	All the time/Often *n* (%)	Sometimes *n* (%)	Rarely/ Not at all *n* (%)
Standardised assessments	24 (92.3)	2 (7.7)	0 (0.0)	19 (73.1)	7 (26.9)	0 (0.0)
Language sampling	17 (65.4)	7 (26.9)	2 (7.7)	11 (42.3)*	9 (34.6)*	6(23.1)*
Criterion-referenced assessments	16 (61.6)	6 (23.1)	4 (15.3)	14 (53.8)	6 (23.1)	6 (23.1)
Dynamic assessments	10 (38.5)	5 (19.2)	11 (42.3)	12 (46.1)	4 (15.4)	10 (38.5)
Processing-dependent measures	2 (7.7)	1 (3.8)	23 (88.5)	2 (7.7)	1 (3.8)	23 (88.5)

Table 4: Standardised assessments used by SLPs when assessing the language skills of Singaporean bilingual children.

Standardised Assessments	Frequency *n* (%)				
	All the time	Often	Sometimes	Rarely	Not at all
Clinical Evaluation of Language Fundamentals USA/UK Editions [CELF] (e.g. *CELF 4:* Semel, Wiig & Secord, 2003)	9 (34.6)	9 (34.6)	6 (23.1)	2 (7.7)	0 (0.0)
Clinical Evaluation of Language Fundamentals Preschool USA/UK Editions [CELF P] (e.g. *CELF P2*: Semel, Wiig & Secord, 2004)	7 (26.9)	8 (30.8)	8 (30.8)	2 (7.7)	1 (3.8)
**Singapore English Action Picture Test [SEAPT]* (Brebner, 2002)	5 (19.2)	8 (30.8)	9 (34.6)	3 (11.5)	1 (3.8)
Preschool Language Scales USA/UK Editions [PLS] (e.g. *PLS 4*: Zimmerman, Pond & Steiner, 2002)	5 (19.2)	7 (26.9)	8 (30.8)	4 (15.4)	2 (7.7)
Peabody Picture Vocabulary Test Editions [PPVT] (e.g. *PVVT 4*: Dunn & Dunn, 2007)	0 (0.0)	2 (7.7)	4 (15.4)	9 (34.6)	11 (42.3)
Expressive Vocabulary Test Editions [EVT] (e.g. EVT2: Williams, 2007)	0 (0.0)	2 (7.7)	4 (15.4)	3 (11.5)	17 (65.4)
Comprehensive Assessment of Spoken Language [CASL] (Carrow-Woolfolk, 1999)	0 (0.0)	1 (3.8)	6 (23.1)	6 (23.1)	13 (50.0)
**Bilingual Language Assessment Battery [BLAB]* (Sze & Rickard-Liow, 2009)	0 (0.0)	1 (3.8)	0 (0.0)	2 (7.7)	23 (88.5)
**Cognitive-Linguistic Profile for Bilingual Preschoolers [CLPBP]* (Lee & Rickard-Liow, 2013)	0 (0.0)	0 (0.0)	1 (3.8)	0 (0.0)	25 (96.2)

*Local standardised assessments

Table 5: Methods used to interpret the performance of Singaporean bilingual children on standardized assessments that were not created and normed for the Singaporean population

Methods	Frequency *n* (%)		
	All the time/ Often	Sometimes	Rarely/ Not at all
Compare a child's performance to normative data provided in assessment manuals	23 (88.4)	3 (11.5)	0 (0.0)
Estimate a child's performance based on clinical experience	21 (80.8)	2 (7.7)	3 (11.5)
Estimate a child's performance using a task analysis checklist approach	13 (50.0)	8 (30.8)	5 (19.2)
Assume that Singaporean children's average language skills are about 6 months behind	8 (30.7)	5 (19.2)	13 (50.0)

Table 6: Challenges in assessment practice when assessing Singaporean bilingual children

Challenges faced	Frequency: All the time/ Often *n* (%)
Availability of 'local' normative data (on standardised assessments)	24 (92.3)
Availability of information on characteristics of language impairment in Singaporean children	24 (92.3)
Availability of developmental data for the Mandarin language	24 (92.3)
Presence of linguistic bias in standardised assessments	24 (92.3)
Availability of developmental data for the Tamil language	23 (88.5)
Availability of developmental data for the Malay language	23 (88.5)
Presence of cultural bias in standardised assessments	23 (88.5)
Availability of locally adapted/developed assessments	21 (80.8)
Differentiating language disorder from language difference	18 (69.2)
Amount of time required to use a full battery of formal and alternative assessments	17 (65.4)
Ability to speak and assess the dominant language of the child	13 (50.0)
Access to interpreters	13 (50.0)
Access to known local standardised assessments	10 (37.8)
Access to international standardised assessments	4 (15.3)

References

American Speech-Language-Hearing Association (2013). *Bilingual Service Delivery.* Retrieved from http://www.asha.org/PRPSpecificTopic.aspx?folderid=8589935225&section=References

Allied Health Professional Council Singapore. (2014). *Report on the Registration of Occupational Therapists, Physiotherapists, Speech-Language Therapists*. Allied Health Professions Council, Ministry of Health, Singapore.

Bedore, L., & Peña, E. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. International *Journal of Bilingual Education and Bilingualism*, *11*(1), 1-29.

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., . . . Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish- English bilingual children. *Bilingualism: Language and Cognition*, *15*(3), 616-629.

Brebner, C. (2010). The development of the Singapore English Action Picture Test: An expressive language screening assessment for Chinese Singaporean preschool children. In M. Cruz-Ferreira (Ed.), Multilingual Norms (pp. 323-341). Bern: Peter Lang Publishing Group.

Brebner, C., Chandler Yeo, H., Goh, M. M., Kam, K. W., & Yeo, W. S. (2015). Tests over time: Evaluating the currency of normative data in a complex multilingual environment. *International Journal of Speech-Language Pathology*, *17*(6), 1-9.

Brebner, C., McCormack, P., & Rickard Liow, S. (2016). Marking of verb tense in the English of preschool English–Mandarin bilingual children: evidence from language development profiles within subgroups on the Singapore English Action Picture Test. *International Journal of Language & Communication Disorders*, *51*(2), 31-43.

Caesar, L. G., & Kohler, P. D. (2007). The State of School- Based Bilingual Assessment: Actual Practice versus Recommended Guidelines. *Language, speech, and hearing services in schools*, *38*(3), 190-200.

Camilleri, B., & Botting, N. (2013). Beyond static assessment of children receptive vocabulary: The dynamic assessment of word learning (DAWL). *International Journal of Language and Communication Disorders*, *48*(5), 565-581.

Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievements. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82-115). New York, NY: Guildford Press.

Carter, J. A., Lees, J. A., Murira, G. M., Gona, J., Neville, B. G. R., & Newton, C. R. J. C. (2005). Issues in the Development of Cross- Cultural Assessments of Speech and Language for Children. *International Journal of Language and Communication Disorders*, *40*(4), 385-401.

Cruz-Ferreira, M., & Ng, B. C. (2010). Assessing multilingual children in multilingual clinics: Insights from Singapore. In M. Cruz-Ferreira (Ed.), *Multilingual Norms* (pp. 343-396). Frankfrut: Peter Lang Publishing Group.

De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language and Communication Disorders*, *46*(6), 613-627.

D'Souza, C., Bird, E. K.-R., & Deacon, H. (2012). Survey of Canadian speech-language pathology service delivery to linguistically diverse clients. *Canadian Journal of Speech-Language Pathology and Audiology*, *36*(1), 18-39.

Gathercole, V. C. M. (2014). Bilingualism matters: One size does not fit all. *International Journal of Behavioral Development*, *38*(4), 359-366.

Gross, M., Buac, M., & Kaushanskaya, M. (2014). Conceptual scoring of receptive and cxpressive vocabulary measures in simultaneous and sequential bilingual children. *American Journal of Speech-Language Pathology*, *23*(4), 574-586.

Guiberson, M., & Atkins, J. (2012). Speech-language pathologists' preparation, practices, and perspectives on serving culturally and linguistically diverse children. *Communication Disorders Quarterly*, *33*(3), 169-180.

Gutiérrez-Clellen, V. F., & Peña, E. (2001). Dynamic assessment of diverse children: A tutorial. *Language Speech and Hearing Services in Schools*, *32*(4), 212-224.

Gutiérrez-clellen, V. F., & Simon-cereijido, G. F. (2009). Using Language Sampling in Clinical Assessments with Bilingual Children: Challenges and Future Directions. *Seminars In Speech And Language*, *30*(04), 234-245.

Hasson, N., Dodd, B., & Botting, N. (2012). Dynamic Assessment of Sentence Structure (DASS): Design and evaluation of a novel procedure for the assessment of syntax in children with language impairments. *International Journal of Language & Communication Disorders*, *47*(3), 285-299.

Hemsley, G., Holm, A., & Dodd, B. (2014). Identifying language difference versus disorder in bilingual children. *Speech, Language and Hearing*, *17*(2), 101-115.

Hoff, E., & Core, C. (2013). Input and language development in bilingually developing children. *Seminars in Speech and Language*, *34*(4), 215.

Hoff, E., & Core, C. (2015). What Clinicians Need to Know about Bilingual Development. *Seminars in Speech and Language*, *36*(2), 89-099.

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech and Hearing Services in Schools*, *43*(1), 81-96.

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders*, *43*(6), 456-473.

Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-based processing tasks separate children with language impairment from typical bilinguals? *Learning Disabilities, Research and Practice*, *21*(1), 19-29.

Laing, S. P., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech and Hearing Services in Schools*, *34*(1), 44-55.

Lidz, C. S. (1991). *Practioner's Guide to Dynamic Assessment*. New York: Guilford.

Okalidou, A., Syrika, A., Beckham, M. E., & Edwards, J. R. (2011). Adapting a receptive vocabulary test for preschool-aged Greek-speaking children. *International Journal of Language and Communication Disorders*, *46*(1), 95-107.

Ooi, C. C.-W., & Wong, A. M.-Y. (2012). Assessing bilingual Chinese– English young children in Malaysia using language sample measures. *International Journal of Speech-Language Pathology*, *14*(6), 499-508.

Paradis, J. (2005). Grammatical morphology in children learning English as a second Language: Implications of similarities with specific language impairment. *Language, Speech and Hearing Services in Schools*, *36*(3), 172-187.

Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, *43*(6), 474-497.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, *43*(1), 93-120.

Peña, E., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech - Language Pathology*, *10*(2), 138-154.

Peña, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., & Sabel, T. (2006). Dynamic assessment of school- age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech, Language, and Hearing Research*, *49*(5), 1037-1057.

Perrie, S., & Core, C. (2006). *Challenges facing monolingual speech-language pathologists in South Florida.* Paper presented at the ASHA Annual Convention, Miami, Florida.

Pieretti, R. A., & Roseberry-McKibbin, C. (2016). Assessment and intervention for English language learners with primary language impairment. *Communication Disorders Quarterly, 37*(2), 117-128.

Roseberry-McKibbin, C., & O'Hanlon, L. (2005). Nonbiased Assessment of English Language Learners: A Tutorial. *Communication Disorders Quarterly*, *26*(3), 178-185.

Royal College of Speech-Language Therapists. (2015). *SLT Assessment and Intervention: Best practice for children and young people in bilingual settings.* Retrieved from https://www.rcslt.org/members/publications/publications2/slt_assessment_and_intervention_main

Singapore Statistics Board. (2010). *Census of Population 2010 Statistical Release 1: Demographic Characteristics, Education, Language and Religion*. Singapore Department of Statistics, Ministry of Trade and Industry, Singapore.

Singapore Statistics Board. (2014). *Population Trends*. Singapore Department of Statistic, Ministry of Trade & Industry, Singapore.

Speech-Language & Audiology Canada. (1997). *SAC Position Paper: Speech-Language Pathology and Audiology in the Multicultural, Multilingual Context.* Retrieved from http://www.sac-oac.ca/professional-resources/resource-library/sac-position-paper-speech-language-pathology-and-audiology

Speech Pathology Australia. (2016). *Position Paper: Working in a Culturally and Linguistically Diverse Society.* Retrieved from https://www.speechpathologyaustralia.org.au/spaweb/Document_Management/Public/Position_Statements.aspx

Stokes, S. F., & Wong, A. M. Y. (1996). Validation of a Cantonese version of the Preschool Language Assessment Instrument. *Asia Pacific Journal of Speech, Language and Hearing*, *1*(2), 75-90.

Stow, C., & Dodd, B. (2003). Providing an equitable service to bilingual children in the UK: a review. *International Journal of Language and Communication Disorders*, *38*, 351-377.

Teoh, W. Q., Brebner, C., & McCormack, P. (2012). Assessing the language skills of children from culturally and linguistically diverse backgrounds: The expressive vocabulary performance of Singaporean English- Mandarin bilingual pre- schoolers. *International Journal of Speech-Language Pathology*, *14*(3), 281-291.

Terrell, S. L., Arensberg, K., & Rosa, M. (1992). Parent-child comparative analysis: A criterion-referenced method for the nondiscriminatory assessment of a child who spoke a relatively uncommon dialect of English. *Language, Speech, and Hearing Services in Schools*, *23*(1), 34-42.

Thordardottir, E. (2015). The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology*, *17*(2), 97-114.

Thordardottir, E., Rothenberg, A., Rivard, M.-E., & Naves, R. (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders*, *4*(1), 1-21.

Thordardottir, E. T. (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. *International Journal of Language and Communication Disorders*, *40*(3), 243.

Williams, C. J., & McLeod, S. (2012). Speech-language pathologists' assessment and intervention practices with multilingual children. *International Journal of Speech-Language Pathology*, *14*(3), 292-305.

Appendix 2: Ethics Approval for Phase One

Dear Wei Qin,

The Chair of the Social and Behavioural Research Ethics Committee (SBREC) at Flinders University considered your response to conditional approval out of session and your project has now been granted final ethics approval. This means that you now have approval to commence your research. Your ethics final approval notice can be found below.

FINAL APPROVAL NOTICE

Project No.:	6691
Project Title:	Assessing the oral language skills of bilingual Singaporean children: Current practices and challenges faced by Speech Language Pathologists in Singapore
Principal Researcher:	Ms Wei Qin Teoh
Email:	teoh0024@flinders.edu.au
Approval Date:	4 November 2014
Ethics Approval Expiry Date:	28 February 2016

The above proposed project has been approved on the basis of the information contained in the application, its attachments and the information subsequently provided with the addition of the following comment:

Additional information required following commencement of research:

1. Please ensure that copies of the correspondence granting permission to conduct the research from the Speech-Language and Hearing Association (SHAS) is submitted to

the Committee on receipt. Please ensure that the SBREC project number is included in the subject line of any permission emails forwarded to the Committee. Please note that data collection should not commence until the researcher has received the relevant permissions (item D8 and Conditional approval response – number 2).

RESPONSIBILITIES OF RESEARCHERS AND SUPERVISORS

1. Participant Documentation

Please note that it is the responsibility of researchers and supervisors, in the case of student projects, to ensure that:

- all participant documents are checked for spelling, grammatical, numbering and formatting errors. The Committee does not accept any responsibility for the above mentioned errors.
- the Flinders University logo is included on all participant documentation (e.g., letters of Introduction, information Sheets, consent forms, debriefing information and questionnaires – with the exception of purchased research tools) and the current Flinders University letterhead is included in the header of all letters of introduction. The Flinders University international logo/letterhead should be used and documentation should contain international dialling codes for all telephone and fax numbers listed for all research to be conducted overseas.
- the SBREC contact details, listed below, are included in the footer of all letters of introduction and information sheets.

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project Number 'INSERT PROJECT No. here following approval'). For more information regarding ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email human.researchethics@flinders.edu.au.

2. Annual Progress / Final Reports

In order to comply with the monitoring requirements of the National Statement on Ethical Conduct in

Human Research (March 2007) an annual progress report must be submitted each year on the 4 November (approval anniversary date) for the duration of the ethics approval using the annual / final report pro forma available from Annual / Final Reports SBREC web page. Please retain this notice for reference when completing annual progress or final reports.

If the project is completed before ethics approval has expired please ensure a final report is submitted immediately. If ethics approval for your project expires please submit either (1) a final report; or (2) an extension of time request and an annual report.

Student Projects

The SBREC recommends that current ethics approval is maintained until a student's thesis has been submitted, reviewed and approved. This is to protect the student in the event that

reviewers recommend some changes that may include the collection of additional participant data.

Your first report is due on 4 November 2015 or on completion of the project, whichever is the earliest.

3. Modifications to Project

Modifications to the project must not proceed until approval has been obtained from the Ethics Committee. Such matters include:

- proposed changes to the research protocol;
- proposed changes to participant recruitment methods;
- amendments to participant documentation and/or research tools; · change of project title;
- extension of ethics approval expiry date; and
- changes to the research team (addition, removals, supervisor changes).

To notify the Committee of any proposed modifications to the project please submit a Modification Request Form to the Executive Officer. Download the form from the website every time a new modification request is submitted to ensure that the most recent form is used. Please note that extension of time requests should be submitted prior to the Ethics Approval Expiry Date listed on this notice.

Change of Contact Details

Please ensure that you notify the Committee if either your mailing or email address changes to ensure that correspondence relating to this project can be sent to you. A modification request is not required to change your contact details.

4. Adverse Events and/or Complaints

Researchers should advise the Executive Officer of the Ethics Committee on 08 8201-3116 or human.researchethics@flinders.edu.au immediately if:

- any complaints regarding the research are received;
- a serious or unexpected adverse event occurs that effects participants;
- an unforseen event occurs that may affect the ethical acceptability of the project.

Appendix 3: Survey Questionnaire: Assessment practices and challenges faced in assessing the oral language skills of Singaporean bilingual children

1) Please indicate the percentage of individuals on your caseload who are: (please enter a figure from 0-100 in each box)

- Pre-schoolers (<ages 6)
- School age (ages 7-12)
- Secondary and above (ages 13-18)
- Adults (ages 18 and above)

2) Please indicate the percentage (estimation) of individuals (ages 12 and below) on your case load who are: (please enter a figure from 0-100 in each box)

- English Dominant
- Mandarin Dominant
- Malay Dominant
- Tamil Dominant
- Others

3) I can speak and provide assessment/intervention in the following languages.

Please choose 'Not at all' for options 'Others 1', 'Others 2' and/or 'Others 3' if they do not apply. If you have selected ‘Others’, a separate field will appear below for you to fill in your responses accordingly.

	Not at all	Minimal	Functional	Proficient
English	○	○	○	○
Mandarin	○	○	○	○

	Not at all	Minimal	Functional	Proficient
Malay	○	○	○	○
Tamil	○	○	○	○
Others 1	○	○	○	○
Others 2	○	○	○	○
Others 3	○	○	○	○

Your responses above indicate that you can speak and provide assessment/intervention in **OTHER** languages, please list the languages accordingly. Please fill in 'NA' if it does not apply.

- Others 1:
- Others 2:
- Others 3:

4) How frequently have you used the following when assessing/profiling the language skills of bilingual Singaporean children (ages 12 and below) **who are ENGLISH DOMINANT**?

Please choose 'Not at all' for options 'Others 1', 'Others 2' and/or 'Others 3' if they do not apply'. If you have selected 'Others', a separate field will appear below for you to fill in your responses accordingly.

	Not at all	Rarely	Sometimes	Often	All the time
Internationally produced language assessment tools (e.g., PLS, CELF)	○	○	○	○	○
Language sample	○	○	○	○	○

	Not at all	Rarely	Sometimes	Often	All the time
Criterion-referenced tests (i.e., Establish baseline - plan intervention goals - document progress)					
Dynamic assessments (i.e., Pretest-Teach-Posttest)					
Processing approaches (e.g., Non-word repetitions, digit span)					
Others 1					
Others 2					
Others 3					

Your responses above indicate that you have used **OTHER** forms/ways of assessing the language skills of bilingual Singaporean children who are English dominant, please list what they are accordingly. Please fill in 'NA' if it does not apply.

- Others 1:
- Others 2:
- Others 3:

5) How frequently have you used the following when assessing/profiling the language skills of bilingual Singaporean children (ages 12 and below) **who are NON-ENGLISH DOMINANT**?

Please choose 'Not at all' for options 'Others 1', 'Others 2' and/or 'Others 3' if they do not apply'. If you have selected 'Others', a separate field will appear below for you to fill in your responses accordingly.

	Not at all	Rarely	Sometimes	Often	All the time
Internationally produced language assessment tools (E.g., PLS, CELF)	◉	○	○	○	○
Translating internationally produced language assessment tools	○	○	○	○	○
Language samples in both languages (English and dominant language)	○	○	○	○	○
Criterion-referenced tests (i.e., Establish baseline - plan intervention goals - document progress)	○	○	○	○	○
Dynamic assessments (i.e., Pretest-Teach-Posttest)	○	○	○	○	○
Processing approaches (e.g., Non-word repetitions, digit span)	○	○	○	○	○
Others 1	○	○	○	○	○
Others 2	○	○	○	○	○
Others 3	○	○	○	○	○

Your responses above indicate that you have used **OTHER** forms/ways of assessing the language skills of bilingual Singaporean children who are Mandarin dominant, please list what they are accordingly. Please fill in 'NA' if it does not apply.

- Others 1:
- Others 2:
- Others 3:

6) How frequently do you use the following **internationally produced language assessment tools** when assessing the language skills of bilingual Singaporean children (ages 12 and below)?

Please choose 'Not at all' for options 'Others 1', 'Others 2' and/or 'Others 3' if they do not apply'. If you have selected 'Others', a separate field will appear below for you to fill in your responses accordingly.

	Not at all	Rarely	Sometimes	Often	All the time
Preschool Language Scale editions:	○	○	○	○	○
Clinical Evaluations of Language Fundamentals Preschool editions:	○	○	○	○	○
Clinical Evaluation of Language Fundamentals editions:	○	○	○	○	○
Comprehensive Assessment of Spoken Language editions:	○	○	○	○	◉
Peabody Picture Vocabulary Test editions:	○	○	○	○	○
Expressive Vocabulary Test editions:	○	○	○	○	○
Others 1:	○	○	○	○	○
Others 2:	○	○	○	○	○
Others 3:	○	○	○	○	○

Your responses above indicate that you have been using **OTHER** internationally produced language assessment tools, please list their names of these assessment tools accordingly. Please fill in 'NA' if it does not apply.

- Others 1:
- Others 2:
- Others 3:

7) How frequently do you use the **following methods** when **profiling the language skills** of bilingual Singaporean children (ages 12 and below) on **internationally produced language assessment tools**?

Please choose 'Not at all' for options 'Others 1', 'Others 2' and/or 'Others 3' if they do not apply'. If you have selected 'Others', a separate field will appear below for you to fill in your responses accordingly.

	Not at all	Rarely	Sometimes	Often	All the time
Compare a child's performance to published/provided normative data in the assessment manuals	○	○	○	○	○
Assume that Singaporean children's average language skills are about 6 months behind published UK/USA/AUS normative data	○	○	○	○	○
Estimate a child's performance using a task/item analysis checklist approach	○	○	○	○	○
Estimate a child's performance (informal) based on clinical experience	○	○	○	○	○
Translate test instructions into child's dominant language (e.g., Mandarin)	○	○	○	○	○
Modify test instructions (e.g., continuing past the test ceiling; not obtaining a baseline)	○	○	○	○	○
Others 1	○	○	○	○	○
Others 2	○	○	○	○	○
Others 3	○	○	○	○	○

Your responses above indicate that you may use **OTHER** methods when profiling the language skills of bilingual Singaporean children on internationally produced language assessment tools, please describe these methods below. Please fill in 'NA' if it does not apply.

- Others 1:
- Others 2:
- Others 3:

8) How frequently have you used the following **locally adapted/created assessment tools**, resources when assessing the language skills of bilingual Singaporean children (ages 12 and below)?

Please choose 'Not at all' for options 'Others 1', 'Others 2' and/or 'Others 3' if they do not apply'. If you have selected 'Others', a separate field will appear below for you to fill in your responses accordingly.

	Not at all	**Rarely**	**Sometimes**	**Often**	**All the time**
Singapore English Action Picture Test (SEAPT) (Brebner, 2002)	○	○	○	○	○
The Singapore Pro-Ed Chart (adapted by Brebner et al., 1996)	○	○	○	○	○
Bilingual Language Assessment Battery (BLAB) (Rickard Liow & Sze, 2009)	○	○	○	○	○
Cognitive-Linguistic Profile for Bilingual Preschoolers (Choo & Rickard Liow, 2013)	○	○	○	○	○
Others 1	○	○	○	○	○
Others 2	○	○	○	○	○
Others 3	○	○	○	○	○

Your responses above indicate that you have used **OTHER** locally adapted/created assessment tools/resources, please list the names of these assessment tools accordingly. Please fill in 'NA' if it does not apply.

- Others 1:
- Others 2:
- Others 3:

Your responses above indicate 'Rarely' or 'Not at all' for some of the above mentioned locally adapted/created assessment tools/resources, please list possible reasons on why so

Answer 

Challenges faced by Speech Language Therapists when assessing multilingual children in Singapore

9) Please indicate the frequency with which you encounter the following **CHALLENGES** when assessing the language skills of bilingual Singaporean children (ages 12 and below).

Please choose 'Not at all' for options 'Others 1', 'Others 2' and/or 'Others 3' if they do not apply. If you have selected 'Others', a separate field will appear below for you to fill in your responses accordingly.

	Not at all	**Rarely**	**Sometimes**	**Often**	**All the time**
Access to internationally produced language assessment tools (e.g., limited access to such tools)	○	○	○	○	○
Access to known locally adapted/produced assessment tools/resources (e.g., limited access to SEAPT, BLAB)	○	○	○	○	○

	Not at all	Rarely	Sometimes	Often	All the time
Availability of locally adapted/produced language assessment tools	○	○	○	○	○
Cultural bias in internationally produced assessment tools (e.g., illustrations of unfamiliar content – campervan; mittens)	○	○	○	○	○
Linguistic bias in internationally produced assessment tools (e.g., Not accepting dialectal variations of Singapore Colloquial English – 'spoil' for 'broken')	○	○	○	○	○
Availability of local normative data to compare a Singaporean child's performance to	○	○	○	○	○
Availability of information on the developmental milestones in the Mandarin language for Singaporean children	○	○	○	○	○
Availability of information on the developmental milestones in the Malay language for Singaporean children	○	○	○	○	○
Availability of information on the developmental milestones in the Tamil language for Singaporean children	○	○	○	○	○
Availability of information on characteristics of language impairment/disorder in Singaporean children	○	○	○	○	◉
Differentiating language disorder/impairment from language difference	○	○	○	○	○

	Not at all	Rarely	Sometimes	Often	All the time
Ability to speak and assess the dominant/primary language of the child	○	○	○	○	○
Amount of time required to use a full battery of formal and informal testing/assessments	○	○	○	○	○
Access to interpreters	○	○	○	○	○
Others 1	○	○	○	○	○
Others 2	○	○	○	○	○
Others 3	○	○	○	○	○

*Your responses above indicate that you have encountered **OTHER CHALLENGES** when assessing the language skills of bilingual Singaporean children, please list them accordingly.

Please fill in 'NA' if it does not apply.

- Others 1
- Others 2
- Others 3

Potential development of local language measures/tools

10) Which of following the assessment measures do you think will be useful to be locally developed to assess the oral language skills of bilingual Singaporean children (ages 12 and below) and for making accurate diagnoses (check as many as are relevant)?

Check any that apply

- ☐ Local standardized assessments
- ☐ Language sampling

- ☐ Dynamic assessment
- ☐ Criterion testing in areas of weakness
- ☐ Processing-based measure

11) What are the factors that you believe need to be considered in assessment development for use with bilingual Singaporean children (ages 12 and below)? (check as many as relevant)?

If you have selected 'Others', a separate field will appear below for you to fill in your responses accordingly.
Check any that apply

- ☐ Test design (e.g., Suitability of test format)
- ☐ Time needed for administration
- ☐ Local normative data
- ☐ Culturally appropriate illustrations
- ☐ Scoring (e.g., Ease of scoring)
- ☐ High reliability and validity of results
- ☐ High sensitivity and specificity (accurately differentiating child with language disorder from typically developing children)
- ☐ Others 1
- ☐ Others 2
- ☐ Others 3

Your responses above indicate **OTHER** factors to be considered in assessments for use with bilingual Singaporean children? please list them accordingly. Please fill in 'NA' if they do not apply.

- Others 1
- Others 2
- Others 3

Please list other any methods that you find essential in assessing the oral language skills of bilingual Singaporean children. Please fill in 'NA' if there are no others.

Answer

Demographic Information

12) I am
Choose one of the following answers

- ○ Singaporean
- ○ Singaporean PR
- ○ Expatriate (Non-local)
- ○ Prefer not to answer

*13) I have been working as ~~an~~ a Speech Language Therapist in Singapore for
Choose one of the following answers

- ○ < 1 year
- ○ 1-3 years
- ○ 4-6 years
- ○ 7-10 years
- ○ 11-15 years
- ○ >15 years
- ○ Prefer not to answer

*14) I currently work in the following settings (please check as many as relevant)
Check any that apply

- ☐ Restructured hospital
- ☐ Private hospital
- ☐ Private clinic/school
- ☐ Community service (e.g., Voluntary Welfare Groups, Government Aided Early Intervention Services)
- ☐ Special school
- ☐ Not working (e.g., post graduate student, in between jobs)
- ☐ Other:

*15) I have worked as a Speech Language Therapist in the following settings (please check as many as relevant)
Check any that apply

- ☐ Restructured hospital
- ☐ Private hospital
- ☐ Private clinic/school
- ☐ Community service (e.g., Voluntary Welfare Groups, Government Aided Early Intervention Services
- ☐ Special school
- ☐ Other:

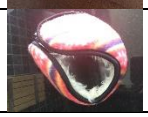
*16) I attended formal training (highest attainment) to be a Speech Language Therapist in
Choose one of the following answers

- ○ Singapore
- ○ Malaysia
- ○ Philippines
- ○ India
- ○ Australia
- ○ New Zealand
- ○ United Kingdom
- ○ Ireland
- ○ USA
- ○ Canada
- ○ Prefer not to answer
- ○ Other:

Appendix 4: DA of Word Learning Skills Scoring Sheet

Participant's code: _____

Test/Retest Phase

Novel words	Unfamiliar Objects (target stimuli)	Test Phase Assessor to say "What is this?" *To ensure the participant is unable to name unfamiliar objects*	Re-test Phase **(Naming)** The participant will be shown each unfamiliar object on its own. Assessor to say "What is this?"	Re-test Phase **(Identification)** All 4 unfamiliar objects to be placed on the table with 4 familiar objects. Assessor to say "Show me _____"
nɛp		1 0	1 0	1 0
hɑn		1 0	1 0	1 0
jɪb		1 0	1 0	1 0
paib		1 0	1 0	1 0
TOTAL				

Teach Phase

Identification

Word (target stimulus)	Number of prompts provided
nɛp	1 2 3 4
hɑn	1 2 3 4
jɪb	1 2 3 4
paib	1 2 3 4
TOTAL	

Naming

Word (target stimulus)	Number of prompts provided
nɛp	1 2 3
hɑn	1 2 3
jɪb	1 2 3
paib	1 2 3
TOTAL	

Teach – IDENTIFICATION

Identification 	***Score*** **word: nɛp**
Look at the things on the table, Teddy asks if you know.. ***Identification Level 1 (context mediation)*** *"Which one is **nɛp**"* The participant does not identify – To provide the next level of mediation	Independent identification Yes (1) No (0)
Identification Level 2 (implicit contextual/language mediation) *"No this is not **nɛp**. Let's try to find the ones you know first. Show me (familiar target 1) and (familiar target 2)" " Now, show me which one is **nɛp**"* The participant does not identify – To provide the next level of mediation	Implicit identification Yes (1) No (0)
Identification Level 3 (explicit context/language mediation) *"No this is not **nɛp**" " This is **nɛp**, this is (familiar target 1), this is (familiar target 2)" "Now, show me which one is **nɛp**"* The participant identifies	Explicit identification Yes (1) No (0) ___/3
Identification 	***Score*** ***word: han***
Look at the things on the table, Teddy asks if you know.. ***Identification Level 1 (context mediation)*** *"Which one is **han**"* The participant does not identify – To provide the next level of mediation	Independent identification Yes (1) No (0)
Identification Level 2 (implicit context/language mediation) *"No this is not **han**. Let's try to find the ones you know first. Show me (familiar target 1) and (familiar target 2)" " Now, show me which one is **han**"* The participant does not identify – To provide the next level of mediation	Implicit identification Yes (1) No (0)
Identification Level 3 (explicit context/language mediation) *"No this is not **han**." " This is **han**, this is (familiar target 1), this is (familiar target 2)" "Now, show me which one is **han**"* The participant identifies	Explicit identification Yes (1) No (0) ___/3

Identification 	***Score*** ***word: jɪb***
Look at the things on the table, Teddy asks if you know.. ***Identification Level 1 (contextual mediation)*** *"Which one is **jɪb**"* The participant does not identify within 8 secs– To provide the next level of mediation ***Identification Level 2 (implicit context/language mediation)*** *"No this is not **jɪb**. Let's try to find the ones you know first. Show me (familiar target 1) and (familiar target 2)" " Now, show me which one is **jɪb**"* The participant does not identify within 8 secs– To provide the next level of mediation ***Identification Level 3 (explicit context/language mediation)*** *"No this is not **jɪb**." This is **jɪb**, this is (familiar target 1), this is (familiar target 2)" "Now, show me which one is _____"* The participant identifies	Independent identification Yes (1) No (0) Implicit identification Yes (1) No (0) Explicit identification Yes (1) No (0) ___/3
Identification 	***Score*** ***word: paib***
Look at the things on the table, Teddy asks if you know.. ***Identification Level 1 (contextual mediation)*** *"Which one is **paib**"* The participant does not identify within 8 secs– To provide the next level of mediation ***Identification Level 2 (implicit context/language mediation)*** *"No, this is not **paib**. Let's try to find the ones you know first. Show me (familiar target 1) and (familiar target 2)" " Now, show me which one is **paib**"* The participant does not identify within 8 secs– To provide the next level of mediation ***Identification Level 3 (explicit context/language mediation)*** *"No, this is not **paib**" " This is **paib**, this is (familiar target 1), this is (familiar target 2)" "Now, show me which one is _____"* The participant identifies	Independent identification Yes (1) No (0) Implicit identification Yes (1) No (0) Explicit identification Yes (1) No (0) ___/3

Teach – NAMING

Naming 	***Score*** **word: nɛp**
Naming level 1 (elicitation question) • *"What is this?"* The participant does not name within 8 secs– To provide the next level of mediation	Independent naming Yes (1) No (0)
Naming level 2 (mediation with semantic prompt) • *"Remember it is a type of chair that Teddy likes to sit" "What is it?"* The participant does not name after 8 secs– To provide the next level of mediation	Naming with semantic prompt Yes (1) No (0)
Naming Level 3 (mediation with phonological prompt) • *"It starts with **/n/**" "What is it?"* The participant does not name after 8 secs – To provide the next level of mediation	Naming with phonological prompt Yes No
Naming Level 4 (mediation with model) • *"This is **nɛp**" "What is it?"*	Naming with model Yes No ___/4

"Wow you just have just remembered and learnt a special name. I can see that you are trying your best to remember the special name for this (hold object). Keep up the good work!" "Now let's look at another one."

Naming 	***Score*** ***word: han***
Naming level 1 (elicitation question) • *"What is this?"* The participant does not name within 8 secs– To provide the next level of mediation	Independent naming Yes (1) No (0)
Naming level 2 (mediation with semantic prompt) • *"Remember it is a type of food that teddy likes to eat" "What is it?"* The participant does not name after 8 secs– To provide the next level of mediation	Naming with semantic prompt Yes (1) No (0)
Naming Level 3 (mediation with phonological prompt) • *"It starts with **/h/**" "What is it?"* The participant does not name after 8 secs – To provide the next level of mediation	Naming with phonological prompt Yes (1) No (0)
Naming Level 4 (mediation with model) • *"This is **han**" "What is it?"*	Naming with model Yes (1) No (0)
	___/4

"Wow you just have just remembered and learnt a special name. I can see that you are trying your best to remember the special name for this (hold object). Keep up the good work!" "Now let's look at another one."

Naming 	***Score*** ***word: jɪb***
Naming level 1 (elicitation question) • *"What is this?"* The participant does not name within 8 secs– To provide the next level of mediation	Independent naming Yes (1) No (0)
Naming level 2 (mediation with semantic prompt) • *"Remember it is a toy that Teddy likes to play" "What is it?"* The participant does not name after 8 secs– To provide the next level of mediation	Naming with semantic prompt Yes (1) No (0)
Naming Level 3 (mediation with phonological prompt) • *"It starts with **/j/**" "What is it?"* The participant does not name after 8 secs – To provide the next level of mediation	Naming with phonological prompt Yes (1) No (0)
Naming Level 4(mediation with model) • *"This is **jɪb**" "What is it?"*	Naming with model Yes (1) No (0) ___/4

"Wow you just have just remembered and learnt a special name. I can see that you are trying your best to remember the special name for this (hold object). Keep up the good work!" "Now let's look at another one."

Naming 	***Score*** ***word: paib***
Naming level 1 (elicitation question) • *"What is this?"* The participant does not name within 8 secs– To provide the next level of mediation	Independent naming Yes (1) No (0)
Naming level 2 (mediation with semantic prompt) • *"Remember it is something that teddy can wear on his ears" "What is it?"* The participant does not name after 8 secs– To provide the next level of mediation	Naming with semantic prompt Yes (1) No (0)
Naming Level 3 (mediation with phonological prompt) • *"It starts with **/p/**" "What is it?"* The participant does not name after 8 secs – To provide the next level of mediation	Naming with phonological prompt Yes (1) No (0)
Naming Level 4 (mediation with model) • *"This is **paib**" "What is it?"*	Naming with model Yes (1) No (0) ___/4

Wow you just have just remembered and learnt a special name for this thing

(Place all target objects on the table)
You worked really hard today!
At first, you didn't know the special names of these.
You were able to (list strategies participant used e.g., listened, looked the object, repeat to yourself) to help yourself to remember the names."

Appendix 5: Ethics Approval for Phase Two (SBREC)

Dear Wei Qin,

The Chair of the Social and Behavioural Research Ethics Committee (SBREC) at Flinders University considered your response to conditional approval out of session and your project has now been granted final ethics approval. This means that you now have approval to commence your research. Your ethics final approval notice can be found below.

FINAL APPROVAL NOTICE

Project No.:	7272
Project Title:	Dynamic assessment of word learning skills (DAWLS): A tool for assessing bilingual children's language skills
Principal Researcher:	Ms Wei Qin Teoh
Email:	teoh0024@flinders.edu.au

Approval Date:	27 June 2016	Ethics Approval Expiry Date:	30 July 2019

The above proposed project has been approved on the basis of the information contained in the application, its attachments and the information subsequently provided with the addition of the following comment(s):

Additional information required following commencement of research:

1. Research Assistants (item B)

As outlined in your response to conditional approval, please ensure that you submit the name and details of any research assistants to be employed via a modification request to the SBREC.

2. Permissions

Please ensure that copies of the correspondence granting permission to conduct the research from

(a) the Ministry of Family and Social Services in Singapore (MSF); and (b) Heads of all schools, kindergartens and childcare centres are submitted to the Committee on receipt. Please ensure that the SBREC project number is included in the subject line of any permission emails forwarded to the Committee. Please note that data collection should not commence until the researcher has received the relevant permissions (item D8 and Conditional approval response – number 13).

3. Other Ethics Committees

Please provide a copy of the ethics approval notice from the Sing-Health Centralised Institutional Review Board (CIRB) on receipt. Please note that data collection should not commence until the researcher has received the relevant ethics committee approvals (item G1 and Conditional approval response – number 14).

RESPONSIBILITIES OF RESEARCHERS AND SUPERVISORS

1. Participant Documentation

Please note that it is the responsibility of researchers and supervisors, in the case of student projects, to ensure that:

- all participant documents are checked for spelling, grammatical, numbering and formatting errors. The Committee does not accept any responsibility for the above mentioned errors.
- the Flinders University logo is included on all participant documentation (e.g., letters of Introduction, information Sheets, consent forms, debriefing information and questionnaires – with the exception of purchased research tools) and the current Flinders University letterhead is included in the header of all letters of introduction. The Flinders University international logo/letterhead should be used and documentation should contain international dialling codes for all telephone and fax numbers listed for all research to be conducted overseas.
- the SBREC contact details, listed below, are included in the footer of all letters of introduction and information sheets.

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project Number 'INSERT PROJECT No. here following approval'). For more information regarding

ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email human.researchethics@flinders.edu.au.

2. Annual Progress / Final Reports

In order to comply with the monitoring requirements of the National Statement on Ethical Conduct in Human Research (March 2007) an annual progress report must be submitted each year on the 27 June (approval anniversary date) for the duration of the ethics approval using the report template available from the Managing Your Ethics Approval SBREC web page. Please retain this notice for reference when completing annual progress or final reports.

If the project is completed before ethics approval has expired please ensure a final report is submitted immediately. If ethics approval for your project expires please submit either (1) a final report; or (2) an extension of time request and an annual report.

Student Projects

The SBREC recommends that current ethics approval is maintained until a student's thesis has been submitted, reviewed and approved. This is to protect the student in the event that reviewers recommend some changes that may include the collection of additional participant data.

Your first report is due on 27 June 2017 or on completion of the project, whichever is the earliest.

3. Modifications to Project

Modifications to the project must not proceed until approval has been obtained from the Ethics Committee. Such proposed changes / modifications include:

· change of project title;

· change to research team (e.g., additions, removals, principal researcher or supervisor change);

· changes to research objectives;

· changes to research protocol;

· changes to participant recruitment methods;

· changes / additions to source(s) of participants;

· changes of procedures used to seek informed consent;

· changes to reimbursements provided to participants;

· changes / additions to information and/or documentation to be provided to potential participants; · changes to research tools (e.g., questionnaire, interview questions, focus group questions);

- extensions of time.

To notify the Committee of any proposed modifications to the project please complete and submit the Modification Request Form which is available from the Managing Your Ethics Approval SBREC web page. Download the form from the website every time a new modification request is submitted to ensure that the most recent form is used. Please note that extension of time requests should be submitted prior to the Ethics Approval Expiry Date listed on this notice.

Change of Contact Details

Please ensure that you notify the Committee if either your mailing or email address changes to ensure that correspondence relating to this project can be sent to you. A modification request is not required to change your contact details.

4. Adverse Events and/or Complaints

Researchers should advise the Executive Officer of the Ethics Committee on 08 8201-3116 or human.researchethics@flinders.edu.au immediately if:

- any complaints regarding the research are received;
- a serious or unexpected adverse event occurs that effects participants;
- an unforeseen event occurs that may affect the ethical acceptability of the project.

Appendix 6: Ethics Approval for Phase Two (Singhealth CIRB)

Next Page

SingHealth
Defining Tomorrow's Medicine

Tel: (65) 6225 0488
Fax: (65) 6557 2464
Singapore Health Services Pte Ltd
31 Third Hospital Avenue
#03-03 Bowyer Block C
Singapore 168753
www.singhealth.com.sg
UEN No 200002698Z

CIRB Ref: **2016/2354**

23 June 2016

Ms Teoh Wei Qin
Department of Child Development
KK Women's and Children's Hospital

Dear Ms Teoh

SINGHEALTH CENTRALISED INSTITUTIONAL REVIEW BOARD (CIRB) APPROVAL

Protocol Title: Dynamic assessment of word learning skills (DAWLS): A tool for assessing Singaporean bilingual children's language skills

We are pleased to inform you that the SingHealth CIRB F has approved the above research project to be conducted in KK Women's and Children's Hospital.

The documents reviewed are:

a) CIRB Application Form dated 7 Jun 2016
b) Participant Information Sheet and Consent Form (Appendix M): Version 2 dated 24 May 2016
c) Child/Participant Assent Form (Appendix N): Version 1 dated 25 Apr 2016
d) Appendix A: Version 1 dated 26 May 2016
e) Appendix B: Version 1 dated 26 May 2016
f) Language Background Questionnaire (Appendix C): English Version 1 dated 24 May 2016
g) Primary Test of Nonverbal Intelligence (PTONI) (Appendix D)
h) Singapore English Action Picture Test (Appendix E)
i) Expressive Vocabulary (EV) (Appendix F)
j) Articulation Screener (Appendix G)
k) Dynamic Assessment of Word Learning Skills (DAWLS) (Appendix H): Version 1 dated 22 Apr 2016
l) Scoring and Data Collection Form (Appendix I): Version 1 dated 22 Apr 2016
m) Letter to School (Appendix K): Version 2 dated 24 May 2016
n) Letter to Parents (Appendix L): Version 2 dated 24 May 2016
o) Summary letter to parent (Appendix O): Version 2 dated 24 May 2016
p) Information Sheet for School (Appendix P): Version 1 dated 24 May 2016

The SingHealth CIRB operates in accordance with the ICH/ Singapore Guideline for Good Clinical Practices, and with the applicable regulatory requirement(s).

The approval period is from **23 June 2016 to 16 May 2017**. The reference number for this study is CIRB Ref: 2016/2354. Please use this reference number for all future correspondence.

The SingHealth CIRB acknowledges receipt of the following translated document:

q) Mandarin Language Background Questionnaire (Appendix C): Version 1 dated 24 May 2016

Please ensure that the translations are an accurate reflection of the original content approved by SingHealth CIRB.

Kindly note that the SingHealth CIRB accepts the authenticity of the translations based on the translation certificates, if any, provided by the Principal Investigator. Consequently, it is the responsibility of the Principal Investigator to ensure that the translations are an accurate reflection of the original approved content.

The following are to be observed upon SingHealth CIRB Approval:

1. No subject should be admitted to the trial before the Health Sciences Authority issues the Clinical Trial Certificate. (only applicable for drug-related studies).

2. The Principal Investigator should ensure that this study is conducted in compliance with the Singapore Guideline for Good Clinical Practice, the ethical guidelines of which are applicable to all studies to be carried out, and to ensure that the study is carried out in accordance to the guidelines and the submitted protocol. The Principal Investigator should meet with his collaborator(s) regularly to assess the progress of the study, and be familiar and comply with all applicable research policies in the Institution.

3. No deviation from, or changes of, the protocol should be initiated without prior written SingHealth CIRB approval of an appropriate amendment, except when necessary to eliminate immediate hazards to the subjects or when the change(s) involve(s) only logistical or administrative aspects of the trial (e.g. change of monitor(s), telephone number(s).

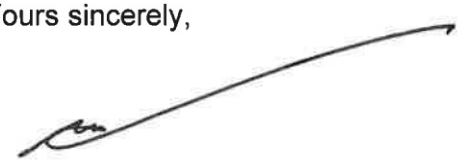
4. Only the approved Participant Information Sheet and Consent Form should be used. It must be signed by each subject prior to enrolling in the study and initiation of any protocol procedures. Two copies of the Informed Consent Form should be signed and dated. Each subject or the subject's legally accepted representative should be given a copy of the signed consent form. The remaining copy should be kept by the PI / medical record.

5. The Principal Investigator should report promptly to the SingHealth CIRB of:
    i. Deviations from, or changes to the protocol including those made to eliminate immediate hazards to the trial subjects.
    ii. Changes increasing the risk to subjects and/or affecting significantly the conduct of the trial.
    iii. All serious adverse events (SAEs) and adverse drug reaction (ADRs) that are both serious and unexpected.
    iv. New information that may affect adversely the safety of the subjects or the conduct of the trial.
    v. Completion of the study.

6. Study Status Report should be submitted to the SingHealth CIRB for the following:
    i. Annual review: Status of the study should be reported to the SingHealth CIRB at least annually using the Study Status Report.

ii. Study renewal: the Study Status Report is to be submitted at least two months prior to the expiry of the approval period. A valid SingHealth CIRB renewal is essential, as any research performed outside of an approved time frame is not legal, and thus not covered by the hospital's research insurance in case of unexpected adverse reactions.
iii. Study completion or termination: the Final Report is to be submitted within three months of study completion or termination.

Yours sincerely,



Dr Aloysius Ho Yew Leng
Chairman
SingHealth Centralised Institutional Review Board F

Enc.

cc: Institution Representative, KKH
Head, Department of Child Development, KKH

Annex 1

LIST OF CIRB F MEMBERS INVOLVED IN THE REVIEW ON 17 MAY 2016			
Name	**CIRB Membership**	**Designation, Institution**	**Gender**
Aloysius Ho Yew Leng	Chairman	Senior Consultant, Haematology, Singapore General Hospital	Male
Francis Chin	Deputy Chairman	Senior Consultant, Radiation Oncology, National Cancer Centre	Male
Jabed Iqbal	Member	Senior Consultant, Pathology, Singapore General Hospital	Male
Edward Poon	Member	Director, Nursing, Ang Mo Kio - Thye Hua Kwan Hospital	Male
Ernest Chew Chin Tiong	Member	Director, St.Luke's Hospital	Male
Colin Phipps Diong	Member	Consultant, Haematology, Singapore General Hospital	Male
Cindy Ng Li Whye	Alternate Member	Principal Physiotherapist, Physiotherapy, Singapore General Hospital	Female

Appendix 7: Participant Information Sheet

PARTICIPANT INFORMATION SHEET

Your child is being invited to participate in a research study.

Before you consent to your child's participation in this research study, please read through the information provided here carefully. If you have any questions about the study, please contact Ms Amanda Loke or Dr Chris Brebner, contact details are provided below.

If you agree to your child's participation in the study, please sign 2 copies of the informed parental consent form (page 6). Please return 1 copy of the of the signed parental consent form in the envelope provided. Please keep 1 copy of the signed parental consent form and this 'Participant Information Sheet' for your reference.

STUDY INFORMATION

Protocol Title:

Dynamic Assessment of Word Learning Skills: A tool for assessing the language skills of Singaporean bilingual children

Investigators and contact details:

Principal Investigator: Teoh Wei Qin
KK Women's and Children's Hospital
Department of Child Development
Email: teoh.wei.qin@kkh.com.sg
Tel: 92706652

Co-Investigator: Amanda Loke
KK Women's and Children's Hospital
Department of Child Development
Email: amanda.loke.hm@kkh.com.sg
Tel: 88097985

Co-Investigator: Dr Sylvia Choo
KK Women's and Children's Hospital
Department of Child Development
Email: sylvia.choo.h.t@singhealth.com.sg

Research Assistant: Enya Mak
KK Women's and Children's Hospital
Department of Child Development
Email: enyamakwt@gmail.com

Co-Investigator: Dr Chris Brebner

Co-Investigator: A/Prof Sue McAllister

Flinders University, Adelaide
Speech Pathology and Audiology
Email: chris.brebner@flinders.edu.au

Flinders University, Adelaide
Speech Pathology and Audiology
Email: sue.mcallister@flinders.edu.au

PURPOSE OF THE RESEARCH STUDY

The research is to help determine if the 'Dynamic Assessment of Word Learning Skills' designed by the investigators (listed above) can be used to assess the language development of Singaporean bilingual children. Information from this study will contribute to the development of local assessment tools to identify Singaporean children who have difficulty learning their language(s).

Your child is invited to participate in this study because your child is attending K1/K2 in a childcare centre/kindergarten. Your child is eligible for this study if he/she meets the following criteria: a) Age 4 years 0 month to 7 years 0 months, b) Chinese race, c) Reported to be exposed to both English and Mandarin language, d) You have no concerns over his/her hearing, speech, and language development OR your child has been referred to receive services under the Development Support Programme.

This study will recruit up to 180 Singaporean bilingual children from childcare centres and kindergarten from July 2016 to July 2017.

STUDY PROCEDURES AND VISIT SCHEDULE

If you agree for your child to take part in this study, there will be tasks that you and your child will need to do:

Parents

You will need to complete a questionnaire titled 'Language Background Questionnaire' as attached.

This questionnaire contains questions about your nationality and educational background, and questions about your child's date of birth, age, schooling and language background. It will take approximately 5-10 minutes to complete.

Child

Your child will be tested on assessments of speech, language and learning skills over 2 sessions. Each session will take between take between 30-60 (maximum) minutes. Each session is designed to fun and engaging with your child. Both sessions will also include play and break time to keep your child motivated and comfortable.

Sessions will take place in your child's school in a corner/room designated by the school. The investigators will arrange the dates and timings of both sessions with the school. Both sessions will only be scheduled during normal school hours and when your child is in school. Prior to each session, your child's agreement to participate will be taken in the presence of a teaching staff. Please be assured that both sessions will be conducted within the line of sight of a teaching staff member.

In session 1, your child will do a:

1. Singapore English Action Picture Test (10-15 mins)

   Your child will be shown 10 pictures and asked to describe what he or she sees in English.

2. Clinical Evaluation of Language Fundamentals Preschool 2 UK version – Expressive Vocabulary subtest (5-10 mins)

   Your child will be shown and asked to name 20 pictures in English.

3. Primary Test of Non-Verbal Intelligence (5-15 mins)

   Your child will be a shown a set of pictures and will be asked to point to the one he/she thinks that does not belong.

In session 2, you child will do a:

1. Articulation Screener (3 mins)

   Your child will be asked to repeat 11 words in English after the investigator.

2. Dynamic Assessment of Word Learning Skills (20-30 mins)

   Your child will be shown 4 unrecognizable objects and be taught on their names (non-real words in English). Your child will be asked to identify and name these objects.

Please take note that the sessions will be audio-recorded so that investigators can analyse your child's spoken responses accurately. The recordings will be stored and locked securely in KK Women's and Children's Hospital with access limited to only investigators listed in this study. Please be assured that any identifiable information (i.e. names) will be removed from the recordings. All softcopies of the audio recordings will be erased after 6 years.

YOUR RESPONSIBILITIES IN THIS STUDY

If you agree for your child to participate in this study, you will:

- Complete and return the Language Background Questionnaire (as attached) with the informed parental consent form (page 6 of this document) in the envelope provided to your child's teacher

If you agree for your child to participate in this study, your child will:

- Take part in 2 sessions of testing on their speech, language and learning skills

WITHDRAWAL FROM STUDY

Participation in this study is entirely voluntary. You are free to withdraw your consent at any point in time. Your child is free to withdraw/decline participation at any time as well. There will be no consequence or effect. Declination/withdrawal will not affect your child's participation in school or/and services that have been offered.

If you decide to withdraw your consent after providing initial consent, please contact and inform co-investigator – Ms Amanda Loke.

The investigators may stop your child's participation in the study at any time for one or more of the following reasons:

- We are matching potential participants for language dominance, gender, age; and may not select your child if there is no match
- Your child demonstrates distress or is unwilling to participate in either session
- Your child's language background details are not complete
- The study is cancelled

WHAT IS EXPERIMENTAL IN THIS STUDY

The study is being conducted because there is a lack of locally-developed assessment measures to evaluate the language learning skills of Singaporean bilingual children.

Your child's participation will contribute to the development of local assessment measures for assessing the language skills of Singaporean bilingual children.

POSSIBLE RISKS, DISCOMFORTS, AND INCONVENIENCES

There is minimal risk to children who participate in this study. The study only selects speech, language and learning measures that are designed for used with preschool children. Your child only needs to provide speaking responses or non-speaking responses (i.e. pointing). These measures have been used in other studies with no reported risk on preschool children.

When your child participates in the study, he/she will be away from classroom activities for 30-60 minutes (maximum) for each session (2 sessions in total). The investigators will coordinate with your child's teacher/school to minimize disruption to your child's classroom activities.

POTENTIAL BENEFITS

If you agree to let your child participate in this study, you may request for a summary of your child's performance on the Singapore English Action Picture Test used in this study without cost. The summary will provide information your child's expressive language skills learning abilities compared to Singaporean children of the same age. If you like a summary letter on your child's performance, please indicate so on the informed consent form (point 7) on page 6. The summary letter will be provided to you in a sealed envelope through the school.

Your child's participation will contribute to the development of local assessment measures for assessing the language skills of Singaporean bilingual children.

ALTERNATIVES

Participation in this study is entirely voluntary. Declination or withdrawal at any point of time will not have any impact on you or your child in any way.

SUBJECT'S RIGHTS

You and your child's participation in this study are entirely voluntary. If you have any questions about the study, you may contact co-investigator Ms Amanda Loke or Dr Chris Brebner (contact information provided below).

By signing and participating in the study, you do not waive any of your legal rights to revoke your consent and withdraw your child from the study at any time.

CONFIDENTIALITY OF STUDY

Information collected for this study will be kept confidential. Your records, to the extent of the applicable laws and regulations, will not be made publicly available. Only Investigators as listed in this study will have access to the confidential information being collected.

However, Regulatory Agencies, Institutional Review Board and Ministry of Health will be granted direct access to check on study procedures and data, without making any of your information public.

By signing the Informed Consent Form attached, you or your legal representative are authorizing collection, access to, use and storage of you and your child's "Personal Data", and (ii) disclosure to authorised service providers and relevant third parties.

"Personal Data" means data about you which makes you and your child identifiable (i) from such data or (ii) from that data and other information which an organisation has or likely to have access.

Research arising in the future, based on this Personal Data, will be subject to review by the relevant institutional review board.

By participating in this research study, you are confirming that you have read, understood and consent to the SingHealth Data Protection Policy - the full version is available at www.singhealth.com.sg/pdpa. Hard copies are also available on request.

Data collected in this study and entered into data collection forms are the property of KK Women's and Children's Hospital and Flinders University. In the event of any publication as a result from this study, you and your child's identity will remain confidential.

COSTS OF PARTICIPATION

There are no costs to your participation. A summary letter of your child's expressive language on the Singapore English Action Picture Test used in the study can be provided on request without cost. You and your child will not be paid for participating in this study.

RESEARCH RELATED INJURY AND COMPENSATION

The Hospital does not make any provisions to compensate study subjects for research-related injury. However, compensation may be considered on a case-by-case basis for unexpected injuries due to non-negligent causes.

By signing this consent form, you will not waive any of your legal rights or release the parties involved in this study from liability for negligence.

WHO TO CONTACT IF YOU HAVE QUESTIONS

If you have questions about this research study, you may contact the following investigators:

- Dr Chris Brebner at chris.brebner@flinders.edu.au
- Ms Amanda Loke at amanda.loke.h.m@kkh.com.sg or at 88097985 (office hours)

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project Number 7722) and SingHealth Centralised Institutional Review Board (Project Number 2016/2354).

If you have questions about your rights as a participant, you can call the SingHealth Centralised Institutional Review Board at 6323 7515 during office hours (8:30 am to 5:30pm). You may also contact Flinders University Social and Behavioural Research Ethics Committee by email at human.researchethics@flinders.edu.au or at +61 8201 3116.

Parental Informed Consent

Details of Research Study

Protocol Title:

Dynamic Assessment of Word Learning Skills: A tool for assessing the language skills of Singaporean bilingual children

Principal Investigator:

Ms Teoh Wei Qin, Department of Child Development, KKH, Tel: 90035586

Child's Particulars

Name: NRIC No.:

Sex: Female/Male Date of birth _____
dd/mm/yyyy

Race: Chinese/ Malay/ Indian /Others (please specify) _____

To be filled by parent / legal guardian / legal representative, where applicable

I, _____ (name of parent/legal guardian) of NRIC/Passport No._____ hereby give consent for my child _____(name) of NRIC/Passport No._____ to participate in the study.

1. I have read the information provided. The nature, risks and benefits of the study have been

   explained clearly to me and I fully understand them.

2. Details of procedures and any risks have been explained to my satisfaction.

3. I am aware that I should retain a copy of the Letter to Parents and Participant Information Sheet.

4. I confirm that I have read, understood and consent to the Singhealth Data Protection Policy.

5. I agree to audio recording of my child's participation.

6. I understand that:
   - My child may not directly benefit from taking part in this study.
   - My child's participation is voluntary. He/she is free to withdraw from the study at any time and is free to decline to answer particular questions.
   - While the information gained in this study may be published, my child will not be identified, and individual information will remain confidential.
   - My child may ask that the audio recording/observation be stopped at any time of the session, and he/she may withdraw at any time from the session without any consequence/effect.
   - Whether my child participates or not, or withdraws after participating, will have no effect on his/her progress in his/her course of study, or results gained.

7. I <u>will like</u>/<u>do not need</u> **(circle one)** a summary letter of my child's performance on the Singapore English Action Picture Test.

_____ _____

Signature of parent /legal guardian Date of signing

参与者信息手册

您的孩子被邀请参与一项研究。

在同意您的孩子参与这项研究之前，请仔细阅读以下的信息。如果您对这项研究有任何的疑问，请向 Amanda Loke 小姐或 Chris Brebner 博士查询。下文将提供联络方式。

如果您同意让您的孩子参与这项研究，请填写并签署两份家长知情同意书（第六页）。请将一份家长知情同意书放入附上的信封并交回。请保留另一份家长知情同意书和这份参与者信息手册作为参考。

研究信息

方案题目：

单字学习能力的动态评量：用于评估新加坡双语儿童的语言能力的工具

研究者及联络方式：

主要研究者：Teoh Wei Qin

竹脚妇幼医院

儿童发展部门

电邮：teoh.wei.qin@kkh.com.sg

电话：92706652

合作研究者：Amanda Loke

竹脚妇幼医院

儿童发展部门

电 邮 ： amanda.loke.hm@kkh.com.sg

电话：88097985

合作研究者：Dr Sylvia Choo

竹脚妇幼医院

儿童发展部门

电 sylvia.choo.h.t@singhealth.com.sg

研究助理：Enya Mak

竹脚妇幼医院

儿童发展部门

电邮：enyamakwt@gmail.com

合作研究者：Chris Brebner 博士

阿德莱德弗林德斯大学

言语病理学和听力学系

合作研究者：Sue McAllister 副教授

阿德莱德弗林德斯大学

言语病理学和听力学系

电邮：chris.brebner@flinders.edu.au　　电邮：sue.mcallister@flinders.edu.au

研究的目的

这项研究将探讨以上研究者所设计的【单字学习能力的动态评量】是否能用于评估新加坡双语儿童的语言发展。这项研究的结果将有助于本地研发评估用具，鉴定有语言学习困难的新加坡儿童。

您的孩子被邀请参与这项研究是因为他/她目前在托儿所/幼稚园就读一/二年级(K1/K2)。如果您的孩子符合下列条件，他/她将有资格参与这项研究：a) 年龄为 4 岁 0 月至 7 岁 0 月，b) 华裔，c) 有接触过英语和华语两种语言，d) 您对于他/她的听力、言语和语言发展没有顾虑*或*您的孩子被转介接受启发补助计划的服务。

这项研究将在 2016 年 7 月至 2017 年 7 月向托儿所和幼稚园招募 180 位新加坡双语儿童。

研究程序和访问行程

如果您同意让您的孩子参与这项研究，您和您的孩子必须执行以下步骤：

家长

您必须填写附件中的语言背景问卷。
这份问卷包含有关您的国籍和教育背景的问题以及有关您孩子的出生日期、年龄、教育背景和语言背景的问题。这份问卷需大约 5 至 10 分钟完成。

儿童

您的孩子将接受两次的评定来测试他/她的语音、语言和学习能力。每次测试的时间长达 30 至 60（最长）分钟。每一次的测试将是生动有趣的。测试中将设有游戏和休息的时间，确保孩子的积极性和舒适性。

每次的测试将在您孩子的学校所设定的角落/课室进行。研究者将与校方安排进行测试的日期和时间。这两次测试只会安排在正常的上课时间以及当您的孩子在学校的时候进行。在进行测试之前，将有教学人员征求您孩子的同意。两次的测试都将在教学人员的视线内进行。

在测试一中，您的孩子将进行：

1. 新加坡英语动作图案测试（10 至 15 分钟）

   研究者将让您的孩子看 10 张图案。他/她必须用英语形容所看到的图案。

2. 英国版本学龄前 2 语言基础的临床评估--表达性词汇测验（5 至 10 分钟）

   研究者将让您的孩子看 20 张图案。他/她必须用英语说出图案的名称。

3. 非文字智力的初试（5 至 15 分钟）

   研究者将让您的孩子看一组图案并请他/她指出他/她认为不属于同一个组别的图案。

在测试二中，您的孩子讲进行：

1. 发音甄别测验(3 分钟)

   研究者将请您的孩子重复研究者所念出的 11 个英文单字。

2. 单字学习能力的动态评量（20 至 30 分钟）

   研究者将让您的孩子看 4 个无法辨认的物件并教他/她这些物件的名称（非真实的英文单字）。研究者将请他/她辨认以及说出物件的名称。

请注意，以上测试的内容将被音频记录，以便让研究者准确地分析您孩子的语音答复。所录制的音频记录将储存以及安全上锁在竹脚妇幼医院内，唯有以上列出的研究者可存取该档案。所有可识别身份的信息（姓名）将在存储档案之前被删除。所有音频记录的电子版将在 6 年后被删除。

您在这项研究中的责任

如果您同意让您的孩子参与这项研究，您必须：

- 将语言背景问卷（附件）和家长知情同意书（这份信函的第 6 页）填写、放入孩子的老师所提供的信封并交回给老师

如果您同意让您的孩子参与这项研究，您的孩子必须：

- 参与两次的言语、语言和学习能力测试

退出研究

参与这项研究是完全自愿的。您可以随时撤销同意。您的孩子也可以随时撤销/拒绝参与。您的决定将没有任何后果。撤销/拒绝参与这项研究将不会影响您孩子在学校的参与或/和其它所提供的服务。

如果您在初步同意后决定撤销同意，请联络并通知合作研究者- Amanda Loke 小姐。

研究者可随时因为一个或多个下列原因而停止您孩子的参与：

- 我们正根据孩子的主导语言、性别和年龄做匹配。如果没有适当匹配的参与者，我们可能不会挑选您的孩子
- 您的孩子在参与任何一次测试时表现出困扰或不愿意参与的意向
- 您的孩子的语言背景资料不完整
- 这项研究被取消了

这项研究中的实验元素

目前新加坡缺乏本地研发并用于评估新加坡双语儿童语言学习能力的评估用具。

您孩子的参与将有助于本地研发评估用具，协助评估新加坡双语儿童的语言能力。

潜在的风险、不适和不便

参与这项研究的儿童风险最低。这项研究只选用为学龄前儿童而设计的言语、语言和学习测度。您的孩子只须作出言语上的反应或非言语的反应（用手指指明）。这些测度已在其它研究中使用，对学龄前儿童没有预知的风险。

当您的孩子参与这项研究时，每一次的测试（总共 2 次）长达 30 至 60 分钟 （最长），而他/她将同时错过课堂活动。研究者们将和您孩子的老师/学校协调如何把对于您孩子的课堂活动所造成的影响降至最低。

潜在获益

如果您的孩子参与这项研究，您可无费用要求领取一份简要介绍您孩子在【新加坡英语动作图案测试】(Singapore English Action Picture Test) 中的成绩的报告。这份报告将为您的孩子的语言表达能力与同龄新加坡儿童的言语表达能力做对比。如果您希望取得一份有关您孩子的成绩的报告，请在第 6 页的知情同意书（第 7 项）上作出说明。我们将通过校方把报告密封在一个信封中并传给你。

您孩子的参与将有助于本地研发评估用具，协助评估新加坡双语儿童的语言能力。

替代方案

参与这项研究是完全自愿的。您在任何时候撤销或拒绝参与这项研究将不会对或您的孩子造成任何影响。

受验者的权利

参与这项研究是完全自愿的。如果您对于这些研究有任何疑问，请联络合作研究者 Amanda Loke 小姐和 Chris Brebner 博士（下文将提供联络方式）。

通过签署和参与这项研究，你不放弃任何撤销同意、以及在任何时候要求孩子退出研究的合法权益。

研究的保密性

这项研究所收集的信息将被严格保密。您的资料，以适用的法律和法规的范围内，不会被公开。唯有以上列出的研究者可存取这些机密资料。

但是，监管机构、机构审查委员会和卫生部将被授权在没有公开您的信息的情况下存取检查研究程序和数据。

通过签署附上的知情同意书，您或您的法律代表将给予收集、获取、使用和储存您和您的孩子的“个人数据”的授权及（ii）向授权的服务提供商和相关第三方公开内容的授权。

“个人数据”是指能够（i）直接或（ii）间接从该数据和某机构所存有或可存取的信息中，识别您和您孩子的数据。

基于这些“个人数据”的未来研究将受到由有关机构审查委员会的审查。

通过参与这项研究，您确认您已阅读、理解并同意新加坡保健数据保护政策--您可浏览 www.singhealth.com.sg/pdpa 阅读完整版。硬拷贝也可应要求提供。

在这项研究中所收集并输入数据收集表格的数据是竹脚妇幼医院和弗林德斯大学的财产。在这项研究后延伸发表任何刊物的情况下，您和您孩子的身份将被严格保密。

参与费用

参与这项研究是免费的。您可无费用要求领取一份简要介绍您孩子在【新加坡英语动作图案测试】(Singapore English Action Picture Test) 中的成绩的报告。参与这项研究，您和您的孩子将不会获得任何补偿。

研究损害赔偿责任

本医院没有任何赔偿受验者研究相关损害的规定，但如果出现非疏忽原因而导致的意外伤害的状况，院方可将依据个别的案例考虑作出赔偿。

通过签署此同意书，你不会放弃任何您的合法权益或释放参与这项研究的各方的疏忽责任。

查询方式

如果您对于这些研究有任何疑问，请联络下列研究者：

- Chris Brebner 博士 chris.brebner@flinders.edu.au
- Amanda Loke 小姐 amanda.loke.h.m@kkh.com.sg 或拨电 88097985(办公时间)

这项研究经由弗林德斯大学社会和行为研究伦理委员会 (研究项目编号 7722) 和新加坡保健集团集中机构审查委员会 (研究项目编号 2016/2354)批准。

如果您对于作为一名参与者的权利有任何疑问，请于办公时间（早上 8 点至下午 5 点 30 分）拨电 6323 7515 向新加坡保健集团集中机构审查委员会查询。您也可通过电邮 human.researchethics@flinders.edu.au 或拨电+61 8201 3116 向弗林德斯大学社会和行为研究伦理委员会查询。

家长知情同意书

研究信息

方案题目：

单字学习能力的动态评量：用于评估新加坡双语儿童的语言能力的工具

主要研究者：

Ms Teoh Wei Qin (张玮芹小姐)，竹脚妇幼医院儿童发展部门，联络号码：92706652

儿童的资料

姓名：　　身份证号码：

性别：女/男　　出生日期：_____

日日/月月/年年年年

种族：华裔/马来裔/印度裔/其他（请注明）　_____

由家长/监护人/法定代表人填写（如适用）

我，_____（家长/监护人姓名），身份证/护照号码_____，在此同意让我的孩子_____（姓名），身份证/护照号码_____，参与这项研究。

8. 我已阅读提供的信息。这项研究的性质、风险和收益已经清楚地向我解说，我也完全的理解。

9. 程序和任何风险详情的解释让我满意。

10. 我知道我应该保留一份家长通知书和参与者信息手册。

11. 我确认已阅读、理解并同意新加坡保健数据保护政策。

12. 我同意我的孩子的参与将被音频记录。

13. 我理解：

- 我的孩子可能无法直接从参与研究中获益。
- 我的孩子的参与是自愿的。他/她可以自由在任何时候退出研究，并可以自由拒绝回答任何的问题。
- 在这项研究中取得的信息可能被刊登，但是我的孩子不会被识别，个人信息也将被严格保密。
- 我的孩子可在被测试的任何时候要求停止音频记录，他/她也能在被测试的任何时候退出，并无须承担任何后果。
- 我的孩子是否参加与否，或参与退出后，对他/她的学习课程没有任何影响。

14. 我希望领取/不需要领取（**请圈其一**）一份有关我的孩子在新加坡英语动作图案测试中的成绩的报告。

_____ _____

家长/监护人签名 日期

Appendix 8: Letter to Parents

LETTER TO PARENTS

Dear Sir/Mdm,

This letter is to introduce Ms Teoh Wei Qin who is undertaking her PhD candidacy in the Department of Speech Pathology and Audiology, School of Medicine, at Flinders University, South Australia. She is also a Senior Speech Language Therapist in the Department of Child Development, KK Women's and Children's Hospital. She will produce her student card and/or staff card, which carries a photograph, as proof of identity.

She is undertaking research leading to the production of a thesis or other publications on the subject of "Dynamic Assessment of Word Learning Skills: A tool for assessing Singaporean bilingual children's language skills". Her supervisors are Dr Chris Brebner and A/Prof Sue McAllister from Flinders University, South Australia.

The research explores if the 'Dynamic Assessment of Word Learning Skills' designed for the purpose of this study can be used to assess the language development of Singaporean bilingual children. Information from this study will contribute to the development of local assessment tools to assist professionals in identifying Singaporean children who have difficulty learning their language(s).

She would like you to assist with the research by consenting your child's participation in the study. If you agree, you will first assist by completing a short questionnaire on your child's language background. Your child will then have their speech, language and learning skills tested on assessments over 2 sessions in school. The assessments will only include speaking activities (i.e. naming, describing pictures) and non-speaking activities (i.e. pointing to picture). Each session will last between 30-60 (max) minutes. Both sessions will also be audio-recorded on a digital audio recorder.

If your child participates in the study, you may request a summary letter of your child's performance on the Singapore English Action Picture Test without cost. The summary will provide information your child's expressive language skills compared to Singaporean children of the same age.

Participation in this study is voluntary. Be assured that any information provided will be treated in the strictest confidence and no children will be individually identifiable in the resulting thesis, report or other publications.

Please refer to and read the 'Participation Information Sheet' where additional details of the study are provided. Any enquiries you may have concerning this project can be directed to me at chris.brebner@flinders.edu.au or Ms Amanda Loke at 88097958 during office hours.

Please indicate your response in the reply slip on the next page. Please return all necessary forms to the school teacher by _____.

Thank you for your attention.

Yours sincerely,

Senior Lecturer & Course Coordinator Master of Speech Pathology, Flinders University

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project Number 7722) and SingHealth Centralised Institutional Review Board (Project Number 2016/2354). If you have questions about your rights as a participant, you can call the SingHealth Centralised Institutional Review Board at 6323 7515 during office hours. You may also contact Flinders University Social and Behavioural Research Ethics committee by email: human.researchethics@flinders.edu.au

REPLY SLIP

Project Title: Dynamic Assessment of Word Learning Skills: A tool for assessing the language skills of Singaporean bilingual children

Principal Investigator: Ms Teoh Wei Qin

Senior Speech Therapist, Department of Child Development, KKH

PhD student, Flinders University, South Australia

Please tick (√)

( ) I would like my child to participate in this study.

Please return the following in the envelope provided:

- **Complete and return this reply slip**. Please keep this letter (page 1) for your own reference
- **Complete and sign 2 copies** of the informed parental consent form (2 copies of page 6 of the 'Participant Information Sheet'). **Please return 1 copy of the signed parental consent form** and keep the 2nd copy for you own reference. Please also keep page 1 to page 5 of the 'Participant Information Sheet' for your own reference
- **Complete and return the Language Background Questionnaire** (as attached)

( ) I wish to decline/withdraw my child's participation in this study.

Please complete and return the following in the envelope provided:

- **Complete and return this reply slip**. Please keep a copy of this letter (page 1) for your own reference

Name of child: _____

Parent's name: _____

Contact details: _____

家长通知书

尊敬的先生/女士，

此书信为您介绍张玮芹小姐 (Ms Teoh Wei Qin)，她目前在南澳弗林德斯大学言语病理学和听力学系攻读博士学位。她也是竹脚妇幼医院儿童发展部门的资深语言治疗师。她将显示附有照片的学生证和/或职员证作为身份证明。

她目前正进行研究并将发表论文或其它刊物，这项研究的题目是“单字学习能力的动态评量：用于评估新加坡双语儿童的语言能力的工具”。她的论文导师是南澳弗林德斯大学的 Chris Brebner 博士 和 Sue McAllister 副教授。

这项研究将探讨专为此研究而设计的【单字学习能力的动态评量】是否能用于评估新加坡双语儿童的语言发展。这项研究的结果将有助于本地研发评估用具，协助专业人员鉴定有语言学习困难的新加坡儿童。

此研究者希望您能协助研究，同意让您的孩子参与这项研究。如果您同意，您首先必须填写一份有关您孩子的语言背景的问卷。您的孩子将在校内接受两次的评定来测试他/她的语音、语言和学习能力。测试的内容仅包括言语活动（命名、形容图案）和非言语活动（指向图案）。每次测试的时间长达 30 至 60 分钟。这两次测试的内容将通过数位录音机被音频记录。

如果您的孩子参与这项研究，您可无费用要求领取一份简要介绍您孩子在【新加坡英语动作图案测试】(Singapore English Action Picture Test) 中的成绩的报告。这份报告将为您的孩子的语言表达能力与同龄新加坡儿童的言语表达能力做对比。

您的孩子参与这项研究是自愿的。您所提供的资讯是保密的，而这项研究后所延伸发表的论文、报告或其它刊物的内容中将不会单独识别任何儿童。

欲知更多有关这项研究的详情，请参考参与者信息手册。若您对这项研究有任何疑问，请通过电邮 chris.brebner@flinders.edu.au 向我查询或于办公时间拨电 88097958 向 Amanda Loke 小姐查询。

请在下一页的回条说明您的回应。请将所需表格在_____之前交回校内的老师。

感谢您的关注。

敬上

弗林德斯大学言语病理学硕士的高级讲师和课程协调员

这项研究经由弗林德斯大学社会和行为研究伦理委员会 (研究项目编号 7722) 和新加坡保健集团集中机构审查委员会 (研究项目编号 2016/2354)批准。如果您对于作为一名参与者的权利有任何疑问，请于办公时间拨电 6323 7515 向新加坡保健集团集中机构审查委员会查询。您也可通过电邮向弗林德斯大学社会和行为研究

回条

研究题目：单字学习能力的动态评量：用于评估新加坡双语儿童的语言能力的工具

主要研究者：张玮芹小姐

竹脚妇幼医院儿童发展部门资深语言治疗师

南澳弗林德斯大学博士研究生

请在适当的括弧内打勾（√）

(　) 我同意让我的孩子参与这项研究。

请将以下文物放入附上的信封并交回：

- **填写并交回这份回条**。请保留这份信函（第 1 页），供自己参考
- **填写并签署两份**家长知情同意书 (参与者信息手册第六页 – 两份)。**请交回一份家长知情同意书，** 请保留第二份供自己参考。请保留参与者信息手册的第 1 至第 5 页，供自己参考
- **填写并交回语言背景问卷**（附件）

(　) 我拒绝/撤回让我的孩子参与这项研究。

请将以下文物放入附上的信封并交回：

- **填写并交回这份回条**。请保留这份信函（第 1 页），供自己参考

孩子的姓名: _____

家长的姓名: _____

联络号码: _____

Appendix 9: Language Background Questionnaire

Language Background Questionnaire

We would appreciate it if you could fill in the blanks or tick the appropriate answer. Thank you!

Your child's details

Name _____ Today's date _____

Date of birth _____ Gender ❒ Male ❒ Female

Number of years in school _____ years

Has your child lived in another country for some time? ❒Yes ❒ No

If yes, how long was that for?_____ years _____ months

Do you have concerns with your child's speech and language development? ❒Yes ❒ No

If yes, please state your concerns:
_____

Do you have concerns with your child's hearing status? ❒Yes ❒ No

If yes, please state your concerns:
_____

Has your child received/is your child receiving any therapy services (i.e. speech therapy) ❒Yes ❒ No

If yes, please state what therapy and when: _____

Parents' details

Mother's name _____

Occupation _____

Highest level of education: ❒ PSLE ❒ O levels ❒ A levels ❒ Poly diploma ❒ University ❒ Others:_____

Are you a Singaporean citizen? ❒Yes ❒ No

If no, what's your nationality?_____ If no, how long have you lived in Singapore? ____ *years*

Father's name _____

Occupation _____

Highest level of education: ❒ PSLE ❒ O levels ❒ A levels ❒ Poly diploma ❒ University ❒ Others:_____

Are you a Singaporean citizen? ❒Yes ❒ No

If no, what's your nationality?_____ If no, how long have you lived in Singapore? ____ *years*

Current family housing: ❒ HDB 1/2 room ❒ HDB 3 rooms ❒HDB 4 rooms ❒ HDB 5 room/Executive

❒ Private (Condominiums/Landed)

Your language use

Please tick the language(s) that you and your husband/wife use with your child, and write down how much of the time you and your husband/wife use this language (e.g., 90% English, 10% Teochew).

Mother's language with child	***Father's language with child***	***Parents' language with each other***
____ % ❒ English	____ % ❒ English	____ % ❒ English
____ % ❒ Mandarin	____ % ❒ Mandarin	____ % ❒ Mandarin
____ % ❒ Dialect _____	____ % ❒ Dialect _____	____ % ❒ Dialect _____
____ % ❒ Malay	____ % ❒ Malay	____ % ❒ Malay
____ % ❒ Tamil	____ % ❒ Tamil	____ % ❒ Tamil
____ % ❒ Others _____	____ % ❒ Others _____	____ % ❒ Others _____

Other caregivers' language with your child

Do you have a maid? ❒ Yes ❒ No

Who looks after your child **most** of the time?

❒ Mother ❒ Father ❒ Maid ❒ Mother's parents ❒ Father's parents ❒ Child-care staff ❒ Baby-sitter

❒ Others (specify) _____

Please tick the language(s) that other main caregivers (e.g., maid or mother's/father's parents) use with your child, and write down how much of the time they use this language with your child (e.g., 60% Teochew, 40% English).

Other caregivers' language		
____ % ❒ English	____ % ❒ Malay	____ % ❒ Others
____ % ❒ Mandarin	____ % ❒ Tamil	____ % ❒ Others

Your child's language proficiency

- Please rank the languages that your child speaks according to how well he/she speaks that language. For rank 1, write the language he/she speaks best; for rank 2, write the next language he speaks best.
- Then write down the age at which he/she was first exposed to this language.
- Finally, please circle a number to rate how good your child's understanding for this language is, and circle a number to rate how good your child's speaking ability for this language is.

Here's a finished example:

Rank	*Language*	*Age of first exposure*		*Not good*			*Average*			*Very good*
			Understanding	*1*	*2*	*3*	*4*	*5*	*6*	*7* (circled)
	English	*From birth*	*Speaking*	*1*	*2*	*3*	*4*	*5*	*6* (circled)	*7*
1 →										
				Not good			*Average*			*Very good*
			Understanding	*1*	*2*	*3*	*4*	*5* (circled)	*6*	*7*
2	*Mandarin*	*3 years*	*Speaking*	*1*	*2*	*3*	*4*	*5* (circled)	*6*	*7*
				Not good			*Average*			*Very good*
			Understanding	*1*	*2*	*3*	*4*	*5* (circled)	*6*	*7*
3	*Teochew*	*From birth*	*Speaking*	*1*	*2*	*3*	*4* (circled)	*5*	*6*	*7*

Please fill in the blanks and circle a number on the rating scale.

Rank	Language	Age of first exposure		Not good			Average			Very good
			Understanding	1	2	3	4	5	6	7
			Speaking	1	2	3	4	5	6	7
1										
				Not good			Average			Very good
			Understanding	1	2	3	4	5	6	7
2			Speaking	1	2	3	4	5	6	7

		Not good			Average		Very good	
	Understanding	1	2	3	4	5	6	7
3	Speaking	1	2	3	4	5	6	7

语言背景问卷调查

请完成下列问题，在适合的地方填写或打勾，谢谢！

您孩子的详细信息

姓名 _____ 填表日期 _____

生日 _____ 性别 ❒ 男 ❒ 女

入学年数 _____ 年

您的孩子曾经有在国外居住吗？ ❒有 ❒没有

如回答有，则居住年限为_____ 年_____ 月

你对您的孩子的语言发长展有什么关注吗？ ❒有 ❒没有

如回答有，请例下你的关注 _____

你对您的孩子的听力状态有什么关注吗？ ❒有 ❒没有

如回答有，请例下你的关注 _____

您的孩子曾经/现在有接受治疗吗 （比如：语言治疗？）❒有 ❒没有

*如回答有，*请例下治疗的详情 和接受治疗的时间段_____

家长信息

母亲姓名 _____职业 _____

最高学历 ❒ 小学 ❒ 中学 ❒高中 ❒ 理工学院 ❒ 大学 ❒其他 _____

新加坡公民 ❒是 ❒不是

如非新加公民，请例下您的国籍_____. 在新加坡居住年限 _____ 年

父亲姓名 _____ 职业 _____

最高学历 ❒ 小学 ❒ 中学 ❒高中 ❒ 理工学院 ❒ 大学 ❒其他 _____

新加坡公民 ❒是 ❒不是

如非新加公民，请例下您的国籍_____. 在新加坡居住年限 _____ 年

目前的家庭住房 ❒ HDB 一房/两房 ❒ HDB 三房 ❒HDB 四房 ❒ HDB 五房以上 ❒ 私人主宰/公寓

语言使用情况

请勾选您和您先生/太太与您孩子交流所使用的语言，并估计该语言所使用的比例（如英语 90%，潮州话 10%）。

妈妈与孩子的交流语言	**爸爸与孩子的交流语言**	**父母之间的交流语言**
____ % ❒ 英语	____ % ❒ 英语	____ % ❒ 英语
____ % ❒ 华语	____ % ❒ 华语	____ % ❒ 华语
____ % ❒ 方言 _____	____ % ❒ 方言 _____	____ % ❒ 方言 _____
____ % ❒ 马来语	____ % ❒ 马来语	____ % ❒ 马来语
____ % ❒ 坦米尔语	____ % ❒ 坦米尔语	____ % ❒ 坦米尔语
____ % ❒ 其他 _____	____ % ❒ 其他 _____	____ % ❒ 其他 _____

育儿人员/女佣与您孩子的交流语言

您是否雇有育儿人员/女佣？ ❒ 是 ❒ 否

谁对您的孩子照顾时间最多？（只选一）

❒ 妈妈 ❒ 爸爸 ❒ 育儿人员/女佣 ❒ 外祖父母 ❒ 祖父母 ❒ 幼托机构职员 ❒ 临时婴孩照顾着

❒ 其他 (请注明) _____

请勾选除父母外其他育儿人员（如祖父母或外祖父母）与孩子的交流语言，并估计该语言所使用的比例（比如，潮州话 60%，英语 40%）。

其他育儿人员所使用语言

____ % ❒ 英语	____ % ❒ 马来语	____ % ❒ 其他
____ % ❒ 华语	____ % ❒ 坦米尔语	____ % ❒ 其他

您孩子的语言流利程度

- 请为您孩子所说语言的流利程度排名。1 表示最流利语言，2 其次流利的语言，以此类推。
- 写下您孩子初次接触到相应语言的年龄。
- 最请圈出您孩子对相应语言的使用能力等级，包括理解能力和表达能力。

例子如下：

排名	*语言*	*初次接触年龄*		*不好*			*一般*			*很好*
			理解能力	*1*	*2*	*3*	*4*	*5*	*6*	*7*
			表达能力	*1*	*2*	*3*	*4*	*5*	*6*	*7*
1	*华语*	*出生*								
			不好				*一般*			*很好*
			理解能力	*1*	*2*	*3*	*4*	*5*	*6*	*7*
2	*英语*	*三岁*	*表达能力*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
			不好				*一般*			*很好*
			理解能力	*1*	*2*	*3*	*4*	*5*	*6*	*7*
3	*潮州话*	*出生*	*表达能力*	*1*	*2*	*3*	*4*	*5*	*6*	*7*

请完成下列表格，并圈出相应语言能力等级

排名	语言	初次接触年龄		不好			一般			很好
			理解能力	1	2	3	4	5	6	7
			表达能力	1	2	3	4	5	6	7
1										
				不好			一般			很好
			理解能力	1	2	3	4	5	6	7
			表达能力	1	2	3	4	5	6	7
2										

		不好			一般			很好
	理解能力	1	2	3	4	5	6	7
3	表达能力	1	2	3	4	5	6	7

Appendix 10: Articulation Screener

Participant's code: _____

Articulation Screener

Directions: Administer all items

Say: "We are going to play a word game now. Say these words after me."

Consonants

Phoneme/Transcription	Words	Child's model
h	**h**igh	
n	**n**ose	
p	**p**ie, ho**p**	
b	**b**ee, tu**b**	
j	**y**awn	

Vowels

Phoneme/Transcription	Words	Child's model
ɑ	f**a**rm	
ɛ	g**i**rl	
ai	sk**y**	
ɪ	b**i**g	

Words adapted from Preschool Language Scale 4 United Kingdom Articulation Screener (2009), The Quick Screener (Bowen, 1996)

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2009). *Preschool Language Scales Edition 4* United Kingdom. San Antonio, TX: PsychCorp.

Bowen, C. (1996). *The Quick Screener* retrieved on 15 June 2016 from www.speech-language-therapy.com