**School of Occupational Therapy, Social Work and Speech Pathology**

**Language Assessment for School-aged Children:
Examining the Evidence and Describing Clinical Practice**

**Deborah Ann Denman**

**This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University**

**June 2019**

## Authors Declaration

To the best of my knowledge and belief this thesis does not contain material previously published by any other person. No other person's work has been used without due acknowledgement. This thesis does not contain material which has been accepted for the award of any other degree or diploma in any university.

The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. Human research ethics approval for this research was received from the Curtin University Human Research Ethics Committee (EC00262), Approval Numbers: HRE2017-0126 and HRE2017-0659.


Deborah Denman

Date: 09/06/2019

**Abstract**

Assessment is an important component in service provision for school-aged children with language disorder. Language interventions are planned based on the results of language assessment, therefore, it is important that the assessment practices speech-language pathologists (SLPs) use are effective in accurately identifying the language needs of children. The research conducted in this thesis contributed to the field of child language assessment by providing new research that enhances professional knowledge in three important areas of need. The first research area pertained to the need for information on the psychometric quality of currently available diagnostic language measures to guide SLPs when selecting language measures for use. The second research area related to the need for consensus on terminology as SLPs use various terminologies interchangeably to describe language assessment practices. The third research area related to the need to develop an understanding of the alignment between current clinical practice and evidence-based practice recommendations. Overall, the objective of this thesis is to identify future actions and research directions that may facilitate future implementation of evidence-based practice recommendations for child language assessment. This thesis comprised of 7 chapters. Chapter 1 provided an introduction and overview of the thesis; Chapters 2 to 6 outline the studies that address the three key research areas and Chapter 7 presented the overall findings and conclusions.

The first research area of research need is targeted in Chapter 2. In this study a systematic review was conducted to examine the psychometric quality of currently available comprehensive norm-referenced spoken language measures for school-aged children. This review employed robust methodology through adherence to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guideline for conducting systematic reviews and use of the COnsensus-based Standards for the selection of health Measurement

Instruments (COSMIN) taxonomy and checklist to rate the methodological quality of included studies. A total of 129 studies relating to 15 different comprehensive norm-referenced spoken language measures were identified and rated for methodological quality. The findings showed that limitations exist with regards to the psychometric evidence of all the evaluated comprehensive norm-referenced spoken language measures. Of the measures evaluated, four were identified as having stronger evidence of psychometric quality compared to others. These four language measures were recommended for use.

The second area of research need is addressed in Chapters 3 and 4. During a three-round Delphi study (reported in Chapter 3), Australian SLPs with expertise in child language assessment were asked to indicate their level of agreement with, and give feedback on a proposed taxonomy for describing language assessment practices for school-aged children. Delphi study participants were also asked to apply the taxonomy by categorising assessment practices presented in case studies. A total of 55 participants completed round one, while 43 and 32 completed rounds two and three respectively. A high level of consensus with the taxonomy was achieved in both rounds one and two, thus establishing expert consensus on terminology for describing child language assessment practices. A lower level of consensus was reached when SLPs were asked to apply the taxonomy terminology to describe assessment practices presented in case studies. To further explore SLPs' perceptions of factors that influenced the consistent application of the taxonomy, semi-structured interviews (reported in Chapter 4) were conducted with 13 of the Delphi study participants. In this study strategies to facilitate consistent application of the taxonomy by SLPs were identified.

Chapters 5 and 6 reported on a study that addressed the third research area of research need. An online survey was used to collect information from Australian SLPs regarding their language assessment practices for school-aged children. The terminology from the taxonomy (Chapter 3) was used to enhance the methodological quality of the survey by facilitating

consistent descriptions of assessment practices by the SLPs who completed the survey. Survey findings are reported across two chapters. Chapter 5 reported on the survey responses from 407 SLPs regarding the regularity with which (i.e., how often) they use different types of language assessments. It was identified that SLPs regularly use language assessments that are described as norm-referenced, de-contextualised and targeted at a clinical context; and less regularly use of assessments described as contextualised, activity-focused, dynamic and targeted at home or school contexts. Regression analysis was used to examine factors that influence the regularity with which different types of assessments were used and results indicated that service agency, Australian state, and SLPs years of working experience influenced the regularity with which some types of assessments were used.

In Chapter 5 SLPs from different service agencies also reported on the main challenges they experience in relation to language assessment and the main sources from which they frequently obtain information on child language assessment. The most frequently reported main challenges related to 'limited assessment materials', 'limited time', 'limited access to training' and 'lack of confidence in assessing children from culturally and linguistically diverse backgrounds'. The most frequently reported source of information on language assessment was 'informal discussion with colleagues'. These findings highlight the importance of considering the influence of contextual factors when examining SLPs implementation of evidence-based assessment practice recommendations.

Chapter 6 reported on the survey responses that 335 of the SLPs provided regarding the specific language measures, assessment procedures and assessment delivery methods they use when assessing the language abilities of school-aged children. SLPs also reported on the assessment domains targeted, purposes of use and reasons for which language measures were chosen for use. Findings identified that although over 130 different language measures are used.  Each SLP used (on average) at total of only seven different language measures when

assessing the last 40 children on their caseload. The most regularly used language measures targeted semantics and syntax rather than social abilities and discourse. SLPs also reported using language measures to assess domains or for purposes for which these measures are not ideally suitable for. Only 6.3% of SLPs reported using information and communication technologies as methods of conducting assessment, however, 29.8% reported that other personnel conduct assessment on their behalf. Collectively, these findings indicated that SLPs may not be making evidence-based decisions when choosing assessments for use.

In the final chapter of this thesis (Chapter 7), findings from the studies conducted across all three research areas are discussed within the context of an implementation science framework. Actions that may support implementation of evidence-based practice recommendations by SLPs are identified and discussed. These actions are important for improving language assessment practice for school-aged children which will, in turn, contribute to improving outcomes children with language disorder.

# Contributors

**Supervisors**

Professor Reinie Cordier

Professor Renée Speyer

Associate Professor Natalie Munro

Dr Jae-Hyun Kim

**Statistician**

Dr Richard Parsons

**Co-authors**

Associate Professor Wendy Pearce, co-author of the manuscript presented in Chapter 2.

Dr Yu-Wei Chen, co-author of the manuscript presented in Chapter 2

Dr. Nathan J. Wilson, co-author of the manuscript presented in Chapter 4

**Funding**

Australian Government Research Training Program Scholarship and Curtin University Top-Up Scholarship

Deborah Denman

Date: 09/06/2019

# Acknowledgements

As with many of life's endeavours, it takes a village to do research. I certainly know that I cannot thank my 'village' enough for assisting me through this project.

Foremost, I extend my sincerest gratitude to my supervisors, Professor Reinie Cordier, Professor Renee Speyer, Associate Professor Natalie Munro and Dr Jae-Hyun Kim. Thank you for setting the bar high and for not ever doubting that I would reach it. Your expertise, kindness and unwavering support have been invaluable and I feel truly fortunate to have been guided by supervisors whose passion for research is both motivating and inspiring.

I feel exceptionally blessed to have my partner and children, whose presence has been the 'perspective' at the end of every long day. Thank you for your never-ending patience throughout this PhD process, for always being my cheer squad and for not letting me forget that I have a sense of humour. I am also grateful for my parents, who raised me to see possibilities rather than limitations. How lucky I am to have all of you in my life.

This project would not exist without the speech pathologists who participated in this study, so thank you all for enthusiastically giving up time in your busy work schedules to be research participants. Thanks also to my friends and colleagues who encouraged me along this research journey – many of you probably do not even know the value that your small words of positivity and encouragement had during those particularly long and tiring weeks of data analysis or writing.

And last, but certainly not least, I acknowledge all the children with language disorders that I have worked with over the years. You are the 'faces on the data'. Enhancing service provision and improving outcomes for you is the important endeavour that has fueled this project.

Thank you, 'Village'.

**Table of Contents**

# List of Tables

## List of Figures

## Explanation of Terms

### Language Disorder

Consistent with the terminology agreed-upon in a recent consensus study by Bishop, Snowling, Thompson, Greenhalgh, and CATALISE-2 consortium (2017); the term 'language disorder' was used in this thesis to refer to any child experiencing difficulties comprehending or producing spoken and written language relative to age expectations, with these difficulties associated with limitations in daily functioning. This definition includes children with developmental language disorder, defined when language difficulties are not explained by the presence of another known diagnosis. It also includes children with language disorder associated with another biomedical condition, such as intellectual disability, Autism Spectrum Disorder or sensory impairment. Current literature does not identify that language assessment practice for school-aged children with language disorder varies (or should vary) depending on the child's diagnosis (Bishop, Snowling, Thompson, & Greenhalgh, 2016; Conti-Ramsden, Simkin, & Botting, 2006; Loucas et al., 2008). Therefore, it is appropriate for a wide variety of conditions to be included in this thesis under the term "language disorder".

### Assessment

In this thesis the term 'assessment' is used broadly to refer to any action that involves collection of information on a child's status, abilities or needs for the purpose of informing service provision (American Educational Research Association, 2014; Paul & Norbury, 2012). The term 'language measure' was used to refer specifically to assessments that have set guidelines for administration and scoring (i.e., standardised assessments). The term 'assessment procedure' was used to refer specifically to assessments that do not have set guidelines for administration and scoring (i.e., non-standardised assessments) (American Educational Research Association, 2014).

**Primary school-aged children**

Although slight variation exists across individual states, children in Australia typically begin primary school between the ages of 4.5-5.5 years and complete primary school between 11.5-12.5 years, depending on birth month. Therefore, in this thesis the term "primary school-aged children" refers to children aged between 4-12 years.

**Publications and Presentations Arising from this Thesis**

**Peer-Reviewed Publications**

Journal Article 1:

Denman, D., Speyer, R., Munro, N., Pearce, W., Chen, Y., & Cordier, R. (2017).

Psychometric properties of language assessments for children aged 4-12 years: A

systematic review. *Frontiers in Psychology, 8*(1515). doi:10.3389/fpsyg.2017.01515

Journal Article 2:

Denman, D., Kim, J.-H., Munro, N., Speyer, R., & Cordier, R. (2019). Describing language

assessments for school-aged children: A Delphi study. *International Journal of*

*Speech-Language Pathology*. Advance online publication.

doi:10.1080/17549507.2018.1552716

Journal Article 3:

Denman, D., Wilson, N. J., Munro, N., Kim, J.-H., Speyer, R., & Cordier, R. (Under Review

with Communication Disorders Quarterly). Factors influencing speech-language

pathologists' application of terminology for describing child language assessments.

Journal Article 4:

Denman, D., Cordier, R., Kim, J.-H., Munro, N., & Speyer, R. (Under Review with Journal

of Communication Disorders). Language assessment practices for primary school

children (Part I): What factors influence assessment use?

Journal Article 5:

Denman, D., Cordier, R., Munro, N., Kim, J.-H., & Speyer, R. (Under Review with Journal

of Communication Disorders). Language assessment practices for primary school

children (Part II): What reasons drive SLP choice?

**Peer-Reviewed Presentations**

Denman, D., Speyer, R., Munro, N., Pearce, W., Chen, Y., & Cordier, R. (2015). *Psychometric properties of language assessments for children aged 4-12 Years: A systematic review*. Paper presented at the Speech Pathology Australia National Conference, Canberra, Australia.

Denman, D., Munro, N., Kim, J.-H., Speyer, R., & Cordier, R. (2018). *Child language assessment and intervention: Development of a taxonomy to describe what we do*. Paper presented at the International Clinical Phonetics and Linguistics Conference, Malta.

Denman, D., Kim, J.-H., Munro, N., Speyer, R., & Cordier, R. (2019). *Child language assessment and intervention: A framework to describe what we do*. Poster presented at the Speech Pathology Australia and New Zealand Speech-Language Therapists' Association Joint Conference, Brisbane, Australia

Denman, D., Munro, N., Kim, J.-H., Speyer, R., & Cordier, R. (2019 ). *Language assessments used for school-aged children: A survey of Australian SLPs*. Poster presented at the Speech Pathology Australia and New Zealand Speech-Language Therapists' Association Joint Conference, Brisbane, Australia.

**Chapter 1.**

**Introduction**

This thesis examines the assessment practices speech language pathologists (SLPs) use for school-aged children and identifies actions for improving clinical assessment practice. In this first chapter, I describe my motivation for researching this topic, relevant background information, and the need for this research. I then conclude this chapter by describing the aims of the project and outline the structure of the thesis.

## 1.1. Motivation for this Research

Prior to undertaking a research doctorate, I worked for over ten years providing speech-language pathology services to school-aged children. During this time, I worked in a range of different service agencies spanning different sectors including disability, education, university and private practice. Children with language disorder always comprised a large proportion of my caseload. These children typically experienced a high degree of functional, personal and educational limitations as a consequence of having language disorder. Concerned about the long-term implications for these children, I sought to choose assessment practices that would most effectively identify the needs of children and thus assist in choosing interventions that would result in optimal outcomes. However, I found this undertaking to be fraught with challenges. I was faced with a large array of possible language assessment options but limited information to guide me in deciding which language measures and assessment procedures to use.

Terms for describing language assessment practices appeared to be used inconsistently across literature, across different service agencies and even within the same service agency. This inconsistency posed a barrier when searching for information on assessment practices and created difficulty when discussing assessment practices with colleagues, as we did not always have a shared understanding of terminology. Lack of terms

for describing assessment practices in detail also made it difficult to identify key similarities and differences between different types of assessments when reflecting on different assessment options.

From observations of clinical practice and discussions with colleagues, it was evident I was not alone in experiencing these challenges. I observed that assessment choice was often directed by SLP personal preference, time constraints or simply by what was available. Assessment practices appeared to be varied and I noticed that my own assessment practices differed depending upon the context in which I was working. Nonetheless, I observed that the focus most often appeared to be on assessing language abilities in discrete-skill tasks, with much less focus on assessing language abilities in everyday communication contexts. These observations lead me to question the alignment between clinical and evidence-based practice and fuelled me with motivation to research language assessment practices for school-age children by undertaking a Doctor of Philosophy. As the following quote reminds us, it is only by continually seeking to improve our clinical practices that we will continue to progress in improving outcomes for the children we serve.

*"Without continual growth and progress, such words as*
*improvement, achievement and success have no meaning"*
– Benjamin Franklin

## 1.2. Children with Language Disorder

Recent research has identified that approximately 10% of school-age children present as having some type of language disorder, which equates to an average of two to three children in every classroom (Norbury et al., 2016). Compared to peers with typical language, children with language disorder are at greater risk of experiencing difficulties with literacy

(Catts, Bridges, Little, & Tomblin, 2008; Fraser & Conti-Ramsden, 2008; Harrison, McLeod, Berthelsen, & Walker, 2009; Mackie, Dockrell, & Lindsay, 2013) and numeracy (Cowan, Donlan, Newton, & Llyod, 2005; Harrison et al., 2009). Studies also report that children with language disorder have increased risks of psychological problems including attention difficulties, social difficulties, depression, low self-esteem, and aggressive behaviour (Liiva & Cleave, 2005; Lindsay, Dockrell, Letchford, & Mackie, 2002; Maggio et al., 2014; McCormack, Harrison, McLeod, & McAllister, 2011; Norbury et al., 2016; Yew & O'Kearney, 2013). These academic and psychological problems have the potential to impact significantly on the quality of life for children with language disorder (Eadie et al., 2018; Feeney, Desha, Khan, & Ziviani, 2017).

Post-school outcomes for individuals with a language disorder are highly variable (Johnson, Beitchman, & Brownlie, 2010), however, as a group, children with language disorder experience poorer psychosocial outcomes compared with peers. A disproportionate number of adolescents and adults with a history of a language disorder have lower levels of employment, poorer quality relationships, and mental health problems (Clarke, Snowling, Truelove, & Hulme, 2010; Conti-Ramsden, Durkin, Simkin, & Knox, 2009; Durkin, Toseeb, Botting, Pickles, & Conti-Ramsden, 2017). There is also a high incidence of previously undiagnosed language disorder amongst adolescents in the justice systems, with studies reporting a language disorder prevalence rate of approximately 60% in this population (Bower et al., 2018; Bryan, Freer, & Furlong, 2007; Lount, Purdy, & Hand, 2017; Snow & Powell, 2008). This myriad of risks highlights the importance of assessment and diagnosis in the primary school years for children with language disorder. Effective assessment of a child's needs is necessary to facilitate the provision of appropriate supports and interventions to assist in potentially averting some of the long-term risks that are associated with language disorder (Norbury et al., 2016).

**1.3. Assessment for Children with Language Disorder**

Language assessment may be conducted using a variety of measures or procedures that assess a child's abilities in different contexts, in different tasks, or through collection different types of data (American Speech and Hearing Association, 2018; Kaderavek, 2015). Terms for describing types of assessments vary across literature (Caesar & Kohler, 2009; Kaderavek, 2015). However, distinctions are often made by categorising assessments as those with set guidelines for administration and scoring and those without set guidelines for administration and scoring (Betz, Eickhoff, & Sullivan, 2013; Caesar & Kohler, 2009).

Assessments with set guidelines for administration and scoring allow for a child's abilities to be assessed in structured tasks in a clinical context. These assessments are often referred to as 'standardised' or 'formal' assessments. Many assessments with set guidelines for administration and scoring also have normative data from an age-matched sample of children and may be called 'norm-referenced' or 'diagnostic' assessments (Kaderavek, 2015). Assessments without set guidelines for administration and scoring allow for a child's abilities to be assessed in natural contexts to provide descriptive data on a child's functional performance. These assessments may be referred to as 'non-standardised', 'informal', 'alternative', 'naturalistic', 'dynamic' or 'functional' (Caesar & Kohler, 2009; Kaderavek, 2015). The term 'language sampling' is also often used to refer to assessments that are 'non-standardised'; however, this term may also be used to refer to any assessment that targets discourse, which subsumes discourse assessments that have set guidelines for administration and yield norm-referenced data (Westerveld & Claessen, 2014).

Language assessment may also be conducted using a number of different methods. For example, assessment may be conducted by SLPs or other personnel such as teachers (Dockrell & Marshall, 2015) and may be conducted either face-to-face or by using information and communication technologies (ICTs; Hodge et al., 2018; Mashima & Doarn,

2009). Assessment data may also be collected through observations of the child's spontaneous talking or by collecting information from parents and teachers regarding the child's current abilities or developmental history (Dockrell & Marshall, 2015).

Data collected from language assessment may be used for a variety of purposes, including identifying the presence of a disorder (Betz et al., 2013) or to determine a child's eligibility for intervention services or educational supports (Dockrell, Lindsay, Letchford, & Mackie, 2006; Fulcher-Rood, Castilla-Earls, & Higginbotham, 2018). Assessment data are also used for the purposes of selecting intervention approaches, identifying appropriate intervention goals, and judging the effectiveness of interventions that are provided (Dockrell & Marshall, 2015; Paul & Norbury, 2012). Given the significant weight these decisions may carry in influencing the services children receive, it is important that SLPs employ evidence-based decision making when selecting assessments for use (Betz et al., 2013; Eadie, 2003).

In Chapter 3 of this thesis, agreement from experts was obtained through a Delphi study on definitions for describing the features and purposes of different language assessments. A table outlining agreed-upon definitions of terms for describing assessments is provided in Supplementary Appendix 3.1.

## 1.4. Psychometric Properties of Language Measures

Norm-referenced assessments are frequently used to make important decisions, such as determining eligibility for intervention services. To ensure accurate decisions are made regarding the need for intervention, norm-referenced language measures chosen for diagnostic use in clinical and research settings should have robust evidence of psychometric quality (Betz et al., 2013). This includes evidence of reliability (internal consistency, test-retest reliability, inter-rater reliability and error measurement), validity (content validity, structural validity, hypothesis testing, cross-cultural validity, criterion validity), responsiveness and interpretability (Mokkink, Prinsen, Bouter, De Vet, & Terwee, 2015). In

addition, norm-referenced measures used for diagnostic purposes should also have evidence for diagnostic accuracy" (Friberg, 2010). Lack of psychometric evidence may mean that the data collected from the assessment is not accurate, thus threatening the soundness of decisions made using the assessment data (Friberg, 2010). Despite this, previous studies have identified limitations in the psychometric quality of available norm-referenced spoken language measures for school-aged children (Andersson, 2005; Betz et al., 2013; Friberg, 2010; McCauley & Swisher, 1984; Plante & Vance, 1994; Spaulding, Plante, & Farinella, 2006). In two previous reviews conducted by Spaulding et al. (2006) and Betz et al. (2013), the manuals from a large number of language measures for children were examined (43 and 55 manuals respectively) for evidence of diagnostic accuracy. Only five out of 43 language measures and 13 out of 55 language measures were identified by the authors as having an acceptable level of diagnostic accuracy. In another review by Friberg (2010), nine language measures identified as having an acceptable level of diagnostic accuracy were reviewed against 11 psychometric criteria, with none of the language measures meeting all 11 criteria. Findings from these reviews identified that limitations may exist with regards to the psychometric properties of norm-referenced language measures and that some language measures may have more robust psychometric evidence than others (Friberg, 2010).

While these reviews were important in highlighting the psychometric limitations of norm-referenced language measures for children, none of these reviews rated the methodological quality of the studies included in the reviews. In addition to considering the outcomes of psychometric studies, the methodology of studies must also be considered. Studies that have not employed appropriate methods for evaluating the psychometric properties of a particular measure cannot be considered as providing robust evidence of psychometric quality. Ratings of methodological quality are needed to determine if the

outcomes reported in studies represent a satisfactory level of evidence for use of a language measure (Marshall, Goldbart, Pickstone, & Roulstone, 2015; Terwee et al., 2012).

Previous reviews examining the psychometric quality of norm-referenced language measures for children were also not systematic in identifying all available assessments and did not include psychometric studies published outside assessment manuals, such as peer reviewed journal articles. Systematic reviews are needed to comprehensively compare all information on available language measures to ensure that all information is considered when making recommendations regarding the language measures that have the best evidence for use (Betz et al., 2013; Terwee et al., 2012).

## 1.5. Evidence-Based Practice Recommendations for Assessment

A lack of consensus exists in relation to assessment and diagnosis of language disorder in children (Reilly et al., 2014), with literature noting variations across political jurisdictions and service agencies with regards to policy related to assessment practice (Ireland, Hall-Mills, & Millikin, 2013; Spaulding, Szulga, & Figueroa, 2012). Although no detailed formal guidelines exist to guide language assessment practice at a national or international level, professional literature does outline a number of evidence-based assessment recommendations for school-aged children with language disorder (Bishop et al., 2016). These recommendations are discussed in this section of the thesis.

**1.5.1. Using assessments that target different contexts and tasks.** Evidence-based practice recommends that SLPs should use a variety of language measures and assessment procedures when evaluating the language abilities of a school-aged child  (American Speech-Language-Hearing Association, 2000; Bishop et al., 2016; Ebert & Pham, 2017; Speech Pathology Australia, 2011a, 2015). Assessment findings may vary depending on the context in which the child's skills are assessed or the types of tasks used within the assessment (Harlaar, DeThorne, Smith, Betancourt, & Petrill, 2016). For example, a number of studies

have identified that results from different language measures conducted in a structured "test-taking" context correlate more highly with each other than they do with results from less structured "language sampling" procedures, suggesting that language performance may be influenced more so by the context than the actual skills being assessed (Dethorne, Johnson, & Loeb, 2005; Harlaar et al., 2016; Ukrainetz & Blomquist, 2002). Similarly, another study by Lennox, Westerveld, and Trembath (2018) identified that performance on a sentence level language measure (Renfrew Action Picture Test; Renfrew, 2010) was not highly predictive of performance on narrative retell task (Profile of Oral Narrative Ability; Westerveld, Gillon, & Boyd, 2012). Therefore, use of assessments that target different contexts and use different types of tasks are needed to fully describe a child's strengths and weaknesses (Bishop et al., 2016). In addition, given that currently available comprehensive norm-referenced spoken language measures have psychometric limitations, assessing language abilities using a range of different assessments reduces the risk of missing important information on a child's language abilities during diagnostic process (Eadie et al., 2014).

**1.5.2. Selecting assessments that match the purpose of the assessment.** Different assessments are suited for different purposes. Language measures with set guidelines for administration and scoring are easily repeatable and relatively quick to administer and analyse (Kaderavek, 2015). However, a weakness of these types of assessments is that tasks are typically not representative of the child's natural communicative environments and are, thus, not suitable for judging functional performance in everyday environments (Kaderavek, 2015; Pavelko, Owens, Ireland, & Hahs-Vaughn, 2016). Many language measures with set guidelines for administration and scoring are designed to cover a comprehensive array of skills in one language measure; however, in doing so may fail to assess each skill in enough depth to identify suitable intervention goals or effectively detect changes over time (Ebert & Scott, 2014).

Assessment procedures without set guidelines for administration and scoring allow for a child's abilities to be evaluated in naturalistic contexts and are thus well-suited for developing intervention goals that target improved performance in everyday activities (Kaderavek, 2015; Yont, Hewitt, & Miccio, 2002). Specific skills may also be examined in greater depth than language measures with set guidelines for administration and scoring might allow. However, assessment procedures without set guidelines for administration and scoring cannot be accurately repeated for the purposes of empirically comparing the performance of different children or comparing the performance of the same child across time (Watson & Pennington, 2015). By using a range of both language measures and assessment procedures, SLPs will have appropriate data to use for the different purposes for which assessment data are often required (Fulcher-Rood et al., 2018).

      **1.5.3. Selecting appropriate assessments for children from CALD backgrounds**. Assessments with norm-referenced data should only be used when the normative sample population matches the child's background (Kaderavek, 2015). When a child's demographics or background is different to the normative population, an accurate comparison of a child's performance in relation to peers is not possible. For this reason, assessments with normative data from monolingual English-speaking children are not appropriate for children who have culturally and linguistically diverse (CALD) backgrounds, including children who are bilingual or learning English as a second language (Arias & Friberg, 2015; Caesar & Kohler, 2007; Pearce & Williams, 2013). For example, one study found the Australian version of the Clinical Evaluation of Language Fundamentals – 4th Edition (CELF-4; Wiig, Semel, & Secord, 2004) to be biased against Aboriginal English language forms and, therefore, likely to identify language difference as being a language disorder in children with Aboriginal backgrounds (Pearce & Williams, 2013). Instead of assessments with norm-referenced data, evidence-based practice guidelines recommend that SLPs should use other types of

assessments to diagnose language disorder in children from CALD backgrounds (Caesar & Kohler, 2007). One recommended alternative is dynamic assessment, which investigates a child's learning potential by examining the child's responses to explicit teaching (Peña, Gillam, & Bedore, 2014). Findings from previous studies have identified that dynamic assessment has clinical utility when differentiating language difference from language disorder in children from CALD backgrounds (Peña, Gillam, & Bedore, 2014; Peña et al., 2006). In addition, dynamic assessment also has usefulness for identifying a child's potential for future language growth (Binger, Kent-Walsh, & King, 2017; Olswang & Bain, 1996).

**1.5.4. Collecting data from parents and teachers as an assessment method**. Evidence-based practice guidelines recommend that SLPs should collect data from parents or teachers when conducting an assessment of a child's language abilities (Bishop et al., 2016). Information from parents and caregivers provide unique information on a child's functional abilities in everyday contexts that supplements results of assessment conducted by SLPs. For example, in a study by Bishop and McDonald (2009) found that identification of language disorder was most accurate when scores from language measures conducted by an SLP were combined with parental report. Similarly, another study measuring intervention outcomes identified that parent report captured information on social abilities that may not be identified through other language measures (Thomas-Stonell, Washington, Oddson, Robertson, & Rosenbaum, 2013). Teacher ratings of a child's language ability also provide valuable diagnostic information, particularly for children from CALD backgrounds, whose abilities may not be accurately assessed through norm-referenced measures (Pearce & Williams, 2013). Collecting information from parents and teachers during the assessment process is also important for developing intervention goals that are aligned with the concerns of families or targeted at addressing educational needs (Crais, Roy, & Free, 2006; Dockrell & Lindsay, 1998).

**1.5.5. Using ICTs as an assessment method**. Research has identified that language assessment for school-aged children may be effectively conducted using ICTs (i.e. telehealth; Edwards, Stredler-Brown, & Houston, 2012)  A study by Waite, Theodoros, Russell, and Cahill (2010b) identified no significant differences between ICT or face-to-face delivery of a norm-referenced language measure for children. In similar studies, a high level of agreement was also obtained between ICT and face-to-face delivery of a range of other language and literacy assessments for children, including measures of phonemic awareness, spelling and reading (Ciccia, Whitford, Krumm, & McNeal, 2011; Waite, Theodoros, Russell, & Cahill, 2010a). Given that many children and families experience lengthy travel times to access services (O'Callaghan, McAllister, & Wilson, 2005; Ruggero, McCabe, Ballard, & Munro, 2012), use of ICTs for conducting language assessment is recommended for improving service accessibility and reducing travel time and costs (Edwards et al., 2012; Mashima & Doarn, 2009).

**1.6. Previous Surveys of Child Language Assessment Practice**

The existence of evidence-based practice recommendations does not guarantee that these recommendations are implemented in clinical practice (Graham et al., 2006). Differences may exist between evidence-based practice recommendations and actual clinical practice (Olswang & Prelock, 2015). In a survey examining SLPs' perceptions of evidence-based practice for children with Autism Spectrum Disorder, 41% of SLPs expressed the view that there are gaps between research evidence and clinical practice (Cheung, Trembath, Arciuli, & Togher, 2013). Other surveys have also identified that although SLPs report valuing evidence-based practice, they experience barriers when implementing evidence-based

practice recommendations into clinical practice (Hoffman, Ireland, Hall-Mills, & Flynn, 2013; Siegel, Maddox, Ogletree, & Westling, 2010; Vallino-Napoli, 2004).

Collecting data on SLPs' language assessment practices is important for determining the extent to which current clinical practice aligns with evidence-based practice recommendations (Caesar & Kohler, 2009; Eadie, 2003). Surveys are well suited for collecting data that allows for large-scale examination of current practice, as this method allows for data to be collected from hundreds of SLPs across different locations (Kelley, Clark, Brown, & Sitzia, 2003). A number of previous surveys have been conducted examining the language assessment practices used by SLPs for school-aged children; however, these surveys have predominantly focused on SLPs in specific service agencies, used single types of assessments or used assessments for specific clinical populations of children. A summary of previous surveys examining SLP language assessment practice is provided in Table 1.1.

Table 1.1

*Summary of Previous Surveys of SLPs Child Language Assessment Practices*

| Author (Year) | Participant country and service agency | Survey sample size | Clinical population targeted and age/grade | [a] Types of assessment practices investigated |
|---|---|---|---|---|
| Arias and Friberg (2015) | SLPs in the USA Employed in schools | 166 | Children from CALD backgrounds Age: 3-21 years | **Standardised:** standardised **Non-standardised:** dynamic, informal, interview with caregiver, language sampling, observations **Other:** N/A |
| Beck (1995) | SLPs in the USA (state of Illinois) Employed in schools | 326 | Any children (no specification) Age: 3-18 years | **Standardised:** names of specific formal assessments used **Non-standardised:** informal observation, language sampling **Other:** N/A |
| Betz et al. (2013) | SLPs in the USA (from multiple states) Employed in schools | 364 | Native English-speaking children with suspected developmental language disorder Age: 5-9 years | **Standardised:** names of specific standardised assessments used for diagnostic purposes **Non-standardised:** N/A **Other:** N/A |
| Caesar and Kohler (2007) | SLPs in the USA (state of Michigan) Employed in schools | 130 | Children from CALD backgrounds Grade: Preschool-high school | **Standardised:** names of specific formal assessments used **Non-standardised:** informal observations, interview with caregiver/teacher, language sampling, observations in class **Other:** N/A |
| Caesar and Kohler (2009) | SLPs in the USA (state of Michigan) Employed in schools | 409 | Native English-speaking children Grade: Preschool-high school age | **Standardised:** names of specific formal assessments used **Non-standardised:** dynamic, informal observations, interview with caregiver/teacher, language sampling, observations in class **Other:** N/A |
| [b] Fulcher-Rood et al. (2018) | SLPs in the USA (from multiple states) Employed in schools | 39 | Children with language disorder Grade: Preschool-high school | **Standardised:** standardised **Non-standardised:** informal **Other:** N/A |
| Gillon et al. (2017) | SLPs from 35 different countries | 1114 | Children with Autism Spectrum Disorder Age not specified | **Standardised**: names of specific assessments used **Non-standardised:** criterion-referenced, dynamic **Other:** N/A |
| Guiberson and Atkins (2012) | SLPs in the USA (state of Colorado) Employed in schools | 154 | Children from CALD backgrounds Age not specified | **Standardised:** standardised **Non-standardised:** N/A **Other:** N/A |
| Huang, Hopkins, and Nippold (1997) | SLPs in the USA (state of Oregon) Employed in schools | 216 | Any children (no specification) Age: Birth-19 years | **Standardised:** standardised **Non-standardised:** N/A **Other:** N/A |
| Hux, Morris-Friehe, and Sanger (1993) | SLPs in the USA (from multiple states) Employed in schools | 239 | Any children (no specification) Grade: Preschool to high school | **Standardised:** N/A **Non-standardised:** language sampling **Other:** N/A |
| Kemp and Klee (1997) | SLPs in the USA (from multiple states) Employed in schools | 253 | Children with language disorder Grade: Preschool | **Standardised:** standardised **Non-standardised:** language sampling **Other:** N/A |

| Author (Year) | Participant country and service agency | Survey sample size | Clinical population targeted and age/grade | [a] Types of assessment practices investigated |
|---|---|---|---|---|
| Pavelko et al. (2016) | SLPs in the USA (from multiple states) Employed in schools | 1336 | Children with suspected language disorder Age: Birth to high school | **Standardised:** N/A<br>**Non-standardised:** language sampling<br>**Other:** N/A |
| [c] Roulstone, Marshall, et al. (2015) | SLPs in the UK Employed in a variety of service agencies | 54 | Children with suspected developmental language disorder Grade: Preschool | **Standardised:** names of specific formal assessments used<br>**Non-standardised:** audio-recording, interview with parent, language sampling, language eliciting activities (books, picture description, posting games), observations, screening assessments devised by SLP, play, questionnaire completed by parent, videoing<br>**Other:** N/A |
| Singh, Chan, and Rusli (2016) | SLPs in Malaysia Employed in a variety of service agencies | 56 | Children with suspected developmental language disorder Age: Approx. 3-5 years | **Standardised:** standardised<br>**Non-standardised:** assessing across different contexts, developmental scales, dynamic, interviews/questionnaires, language sampling, observations<br>**Other:** N/A |
| Teoh, Brebner, and McAllister (2017) | SLPs in Singapore Employed in a variety of service agencies | 26 | Children from CALD backgrounds Age not specified | **Standardised:** standardised<br>**Non-standardised:** criterion-referenced, dynamic, language sampling, processing-dependent measures<br>**Other:** N/A |
| Tucker (2012) | SLPs in the USA (one state only – name of state unspecified) Employed in schools | 170 | Any children (no specification) Grade: Primary to high school | **Standardised:** N/A<br>**Non-standardised:** N/A<br>**Other:** Use of ICTs for conducting screening or assessment (including screening and assessment of language abilities) |
| Watson and Pennington (2015) | SLPs in the UK Employed in a variety of service agencies | 265 | Children and young people with diagnosis of Cerebral Palsy Age: Birth to young adult | **Standardised:** names of specific standardised assessments used<br>**Non-standardised:** observation, assessments devised by SLP<br>**Other:** N/A |
| Westerveld and Claessen (2014) | SLPs in Australia (from multiple states) Employed in a variety of service agencies | 257 | Children with suspected language disorder Age: Birth to high school | **Standardised:** names of standardised language sampling assessments used<br>**Non-standardised:** informal language sampling<br>**Other:** N/A |
| Williams and McLeod (2012) | SLPs in Australia (from multiple states) Employed in a variety of service agencies | [d] 118 | Children from CALD backgrounds Age not specified | **Standardised:** standardised assessments in English, standardised assessments in other languages<br>**Non-standardised:** dynamic, informal, locally developed tests, processing approaches<br>**Other:** N/A |
| Wilson, Blackmon, Hall, and Elcholtz (1991) | SLPs in the USA (state of California) Employed in schools | 266 | Any children (no specification) Age: 4 to 9 years | **Standardised:** names of specific formal assessments used<br>**Non-standardised:** informal assessments devised by SLP<br>**Other:** N/A |

*Note.* [a] The terms listed under the headings standardised and non-standardised follow (where possible) the actual terms used by the authors of each survey to describe types of assessments in their study; [b] In this study, participants provided data through interviews; [c] In this study, 29 participants provided data via focus groups; [d] Number of survey participants who reported on assessment practice; CALD = Culturally and Linguistically Diverse; USA = United States of America; N/A = Not applicable as this survey did not investigate this type of assessment.

The majority of previous surveys have been conducted in the United States of America and focus on the assessment practices of SLPs employed in public schools (Arias & Friberg, 2015; Beck, 1995; Betz et al., 2013; Caesar & Kohler, 2009; Guiberson & Atkins, 2012; Huang et al., 1997; Hux et al., 1993; Pavelko et al., 2016; Tucker, 2012; Wilson et al., 1991). Other surveys have specifically targeted assessment practices for children from (CALD) backgrounds (Arias & Friberg, 2015; Guiberson & Atkins, 2012; Teoh, Brebner, & McCormack, 2012; Williams & McLeod, 2012) or children with specific disabilities (Gillon et al., 2017; Watson & Pennington, 2015). Similarly, some earlier surveys deliberately only targeted single assessments, such as norm-referenced language measures or assessment procedures described as language sampling (Betz et al., 2013; Guiberson & Atkins, 2012; Huang et al., 1997; Hux et al., 1993; Pavelko et al., 2016; Westerveld & Claessen, 2014). Only two previous surveys have collected data from SLPs across different service agencies regarding the range of language assessment practices they use for a broad population of children (Roulstone, Marshall, et al., 2015; Singh et al., 2016). These two surveys involved small samples of SLPs and specifically targeted preschool-aged children (under 6 years). Therefore, further surveys are needed to examine if the practice trends identified in these surveys are representative of SLPs' language assessment practice for school-aged children across a broad context (Fulcher-Rood et al., 2018).

Findings from previous surveys have identified that SLPs use both language measures with set guidelines for administration and scoring and assessment procedures without set guidelines for administration and scoring when assessing the language abilities of school-aged children (Arias & Friberg, 2015; Caesar & Kohler, 2009; Gillon et al., 2017; Singh et al., 2016). However, an over-reliance on language measures with norm-referenced data appears to exist (Fulcher-Rood et al., 2018), particularly when assessing children from CALD backgrounds (Caesar & Kohler, 2007; Fulcher-Rood et al., 2018; Teoh et al., 2017).

Furthermore, some assessments, such as dynamic assessments, appear to be used much less regularly than other types of assessments (Caesar & Kohler, 2009; Teoh et al., 2017; Williams & McLeod, 2012). These findings suggest that SLPs may not always be regularly collecting data from a range of different types of assessments when assessing children with language disorder, however, further information is needed to substantiate this.

Previous survey findings have also identified that SLPs may not be heeding evidence-based practice recommendations when choosing assessments for use (Betz et al., 2013; Fulcher-Rood et al., 2018). For example, studies have identified that SLPs may use norm-referenced language measures for purposes other than what they were designed for, such as planning intervention goals (Beck, 1995; Fulcher-Rood et al., 2018; Huang et al., 1997; Kemp & Klee, 1997). Other studies have identified that SLPs may not be considering available psychometric evidence when choosing language measures for diagnostic purposes (Betz et al., 2013; Fulcher-Rood et al., 2018). Only one previous survey has examined the domains assessed, purposes for use or reasons why SLPs choose specific norm-referenced language measures (Roulstone, Marshall, et al., 2015). In this previous survey, SLPs reported using language measures for assessing domains or for purposes other than what the measure was developed to assess, which may lead to inappropriate conclusions regarding a child's performance. For example, the Renfrew Action Picture Test (RAPT) was reported by SLPs as being used to assess functional communication or to screen language abilities, despite limited evidence that this assessment is suitable for these purposes (Roulstone, Marshall, et al., 2015). However, as this survey is a single study consisting of a small sample of SLPs, more evidence is needed to develop a greater understanding of SLPs use of child language assessments in clinical practice.

There is also a need for further studies examining SLPs use of different assessment delivery methods, such as having other personnel conduct assessments, use of ICTs to

conduct assessments or collecting information from parents and teachers. Previous surveys have identified that SLPs do collect information from parents and teachers as part of their assessment process (Caesar & Kohler, 2007, 2009; Singh et al., 2016), however, these surveys did not examine the types of assessments SLPs use to do this. In addition, although previous studies have examined SLPs attitudes towards service provision via ICTs (Dunkley, Pattie, Wilson, & McAllister, 2010; Hines, Lincoln, Ramsden, Martinovich, & Fairweather, 2015); only one previous survey has examined SLPs use of ICTs to conduct child language assessment (Tucker, 2012). This survey targeted SLPs employed in schools in a single state in the United States of America and was not specifically focused on language assessment practice, therefore further studies are needed to determine the regularity with which the board population of SLPs use ICTs when conducting child language assessment. The practice of having other personnel administer language assessments has also not been previously examined. Obtaining information on the assessment delivery methods SLPs currently use when assessing school-aged children will improve understanding of current speech language pathology language assessment practice and assist in identifying future research directions.

## 1.7. Implementation of Evidence-Based Practice Recommendations

The field of implementation science is concerned with the investigation of the processes involved in the transfer of research evidence into clinical practice (Graham et al., 2006). These processes are acknowledged to be complex and influenced by multiple factors (Straus, Tetroe, & Graham, 2009). Nonetheless, examining the factors that influence the implementation of evidence-based practice recommendations allows for specific strategies to be identified and enacted to promote the successful implementation of evidence-based practice recommendations (Hakkennes & Dodd, 2008). Although literature identifies factors that may act as barriers or facilitators to the successful implementation of evidence-based practice recommendations by health practitioners in general, there is a paucity of research

examining the factors that may influence implementation of evidence-based practice recommendations by SLPs more specifically (Miao, Power, & O'Halloran, 2015). Furthermore, the majority of studies conducted in the speech language pathology field have focused on engagement in evidence-based research activities (Cheung et al., 2013; O'Connor & Pettigrew, 2009; Vallino-Napoli, 2004) or implementation of clinical practice guidelines for adult rehabilitation interventions (Miao et al., 2015; Shrubsole, Worrall, Power, & O'Connor, 2018; Young, Shrubsole, Worrall, & Power, 2018). Therefore, significant gaps in knowledge exist with regards to understanding the factors that may influence implementation of evidence-based practice recommendations for child language assessment. As it is important that evidence-based practice recommendations are implemented successfully into clinical practice, research in this area is a priority (Fulcher-Rood et al., 2018; Olswang & Prelock, 2015; Roulstone, 2001).

One tool designed to help identify factors that may influence health practitioners' implementation of evidence-based practices is the comprehensive, integrated checklist of determinants of practice (the TICD checklist; Flottorp et al., 2013). Developed through a systematic review, the TICD checklist was created from an amalgamation of 12 previously published implementation science frameworks. As such, the TICD checklist provides a comprehensive list of factors to consider when reflecting upon the specific factors that may influence implementation of a specific evidence-based practice recommendation (Flottorp et al., 2013). In this thesis, the factors listed in the TICD checklist are presented diagrammatically with specific application to speech language pathology assessment practice. See Figure 1.1.
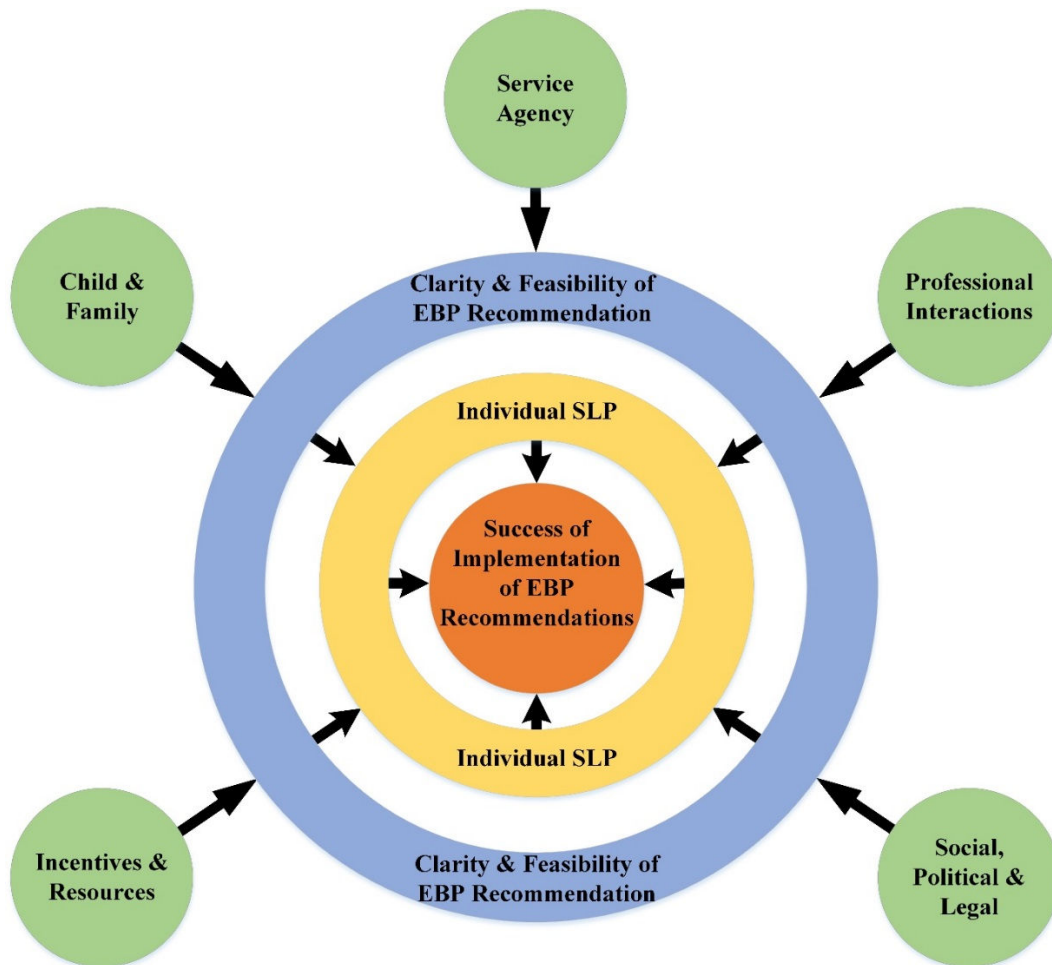
*Figure 1.1.* Factors that may influence SLPs implementation of evidence-based practice recommendations (based on the TICD checklist by Flottorp et al., 2013).

The success with which evidence-based recommendations are implemented is represented in the inner (orange-coloured) circle of Figure 1.1. As individual SLPs are the agents who directly implement recommendations, factors internal to individual SLPs are represented as the first inner (yellow-coloured) ring in this figure. Factors related to the individual SLP include: awareness of and familiarity with recommendation; attitudes and beliefs towards recommendation; and knowledge, skills and confidence of SLPs (Flottorp et al., 2013; Michie et al., 2005).

The clarity and feasibility of the recommendation itself may contribute to the success with which evidence-based practice recommendations are implemented. This is shown as the second outer (blue-coloured) ring in Figure 1.1. These factors include: the quality and

credibility of evidence behind recommendation, clarity and accessibility of recommendation, consistency of recommendation with other recommendations and the feasibility with which the recommendation can be implemented in clinical settings (Cabana et al., 1999; Flottorp et al., 2013).

Factors external to the SLP may also influence implementation of evidence-based practice recommendations (Cabana et al., 1999). These factors are represented by the outer (green-coloured) circles in the Figure 1.1. The 'service agency' circle refers to factors such as service agency regulations or policies, priority given to practice change, capability of leadership and authority and accountability structures (Belkhodja, Amara, Landry, & Ouimet, 2007; Flottorp et al., 2013). The 'child and family' circle refers to factors related to SLPs' perceptions of expectations, preferences and motivation of children and families or actual expectations, preferences and motivations of children and families (Cabana et al., 1999; Flottorp et al., 2013). The 'professional interactions' circle includes factors such as the influence of professional communications and the ability for professional teams to work together (Dopson, FitzGerald, Ferlie, Gabbay, & Locock, 2002; Flottorp et al., 2013). 'Resources and incentives' relates to factors such as availability of required time, materials, technology and professional support or presence of financial or nonfinancial incentives (Flottorp et al., 2013; Harding, 2014). 'Social, political or legal' factors include the influence of funding policies, opinions of influential people, and ethical or legal issues (Flottorp et al., 2013).

In the following section of this thesis, the factors represented in the Figure 1.1 are used to structure a summary of current literature and highlight the knowledge gaps that exist in relation to understanding the factors that influence SLPs implementation of evidence-based language assessment practices for school-aged children.

**1.7.1. Factors related to the individual SLP.** Factors internal to individual SLPs may influence implementation of evidence-based practice recommendations (Cabana et al., 1999; Flottorp et al., 2013). Previous studies investigating the practices of SLPs in adult rehabilitation settings reported that SLPs may lack awareness and familiarity with practice recommendations (Shrubsole et al., 2018). Other studies have identified that, as a whole, SLPs do not frequently access peer reviewed journal articles to source information on evidence-based practice but, instead, rely on information from peers (Hoffman et al., 2013; Vallino-Napoli, 2004). Low use of journal articles for sourcing information specific to child language assessment practice has been reported in earlier surveys of child language assessment practice (Beck, 1995; Wilson et al., 1991); however, as computer and internet access has likely increased since these earlier studies were conducted, further investigation is needed to identify the sources from which SLPs currently obtain information specific to child language assessment practice. This information may assist in understanding the context that surrounds current practice and will help identify future avenues for effectively disseminating information to SLPs on language assessment practice.

Studies examining the practices of SLPs employed in adult rehabilitation settings have also reported that SLPs' motivation to implement evidence-based practice recommendations may be influenced by personal commitment to evidence-based practice or beliefs about the effectiveness of practice recommendations (Miao et al., 2015; Shrubsole et al., 2018). It is also possible that SLPs make assumptions regarding the amount of time or effort required to implement evidence-based practice recommendations, which influences their clinical practice (Michie et al., 2005). To understand the factors that influence implementation of evidence-based practice recommendations for school-aged language assessment specifically, further information is needed with regards to the challenges SLPs perceive in relation to school-aged language assessment.

No studies have specifically examined SLPs' qualifications in relation to child language assessment practice, however, some studies have examined the influence of SLPs' qualifications on evidence-based practice more generally. One previous study found that possession of post graduate qualifications positively influences SLPs' engagement in research activities (Finch, Cornwell, Ward, & McPhail, 2013), however, another study did not identify SLPs' qualifications as a factor that influences application of research findings into clinical practice (Vallino-Napoli, 2004).

A number of previous studies have examined child language assessment practice in relation to SLPs years of working experience (Caesar & Kohler, 2007; Hux et al., 1993; Pavelko et al., 2016; Roulstone, Marshall, et al., 2015). While these studies did not identify years of working experience as a factor that directly influences the types of assessment practices SLPs use, it was found that SLPs with more years of working experience were more likely to make their own assessment protocols, assess in different contexts or rely more on their own judgement when analysing assessment results (Caesar & Kohler, 2007; Pavelko et al., 2016; Roulstone, Marshall, et al., 2015; Wilson et al., 1991). These mixed findings suggest that further investigation is warranted to understand in greater depth the influence of SLPs working experience and qualifications on assessment practice.

SLPs knowledge and training may influence implementation of evidence-based practice recommendations (Flottorp et al., 2013). In previous surveys investigating child language assessment practice, SLPs have reported lack of familiarity and training as barriers to the use of some types of assessments, such as dynamic assessment or language sampling (Arias & Friberg, 2015; Pavelko et al., 2016; Westerveld & Claessen, 2014). Furthermore, when practice recommendations are viewed as challenging to implement, SLPs have reported a lack of self-efficacy in their own ability to change their practice (Kritikos, 2003). However, as previous studies have only examined SLPs knowledge and self-efficacy in relation to

specific assessment contexts or specific population of children (Kritikos, 2003; Pavelko et al., 2016; Santhanam & Parveen, 2018; Westerveld & Claessen, 2014), further evidence is needed to determine the extent to which the broad population of SLPs identify lack of training and knowledge as a challenge to conducting language assessment.

**1.7.2. Factors related to clarity and feasibility of practice recommendations.** Implementation of evidence-based practice recommendations may be influenced by factors related to the practice recommendations themselves (Cabana et al., 1999; Flottorp et al., 2013). A study of SLPs in adult rehabilitation settings reported that successful implementation of evidence-based recommendations may be facilitated by explicitly building SLPs' understanding of how recommendations apply to their specific clinical settings (Miao et al., 2015). However, in a previous survey in the United States of America, 89% of SLPs employed in schools reported not having specific procedural guidelines in their workplaces to guide evidence-based practice (Hoffman et al., 2013). Identifying specific areas of speech language pathology clinical assessment practice that are not well-aligned with evidence-based recommendations may assist in identifying areas where specific clinical practice guidelines may be needed to support clinical practice.

**1.7.3. Factors external to the individual SLP that influence implementation.** Factors in the environment around the SLP may influence implementation of evidence-based practice recommendations (Cabana et al., 1999). These include influences related to 'service agency', 'child and family', 'professional interactions', 'resources and incentives' and 'social, political and legal' factors.

*1.7.3.1. Service agency.* Previous studies have reported that service agency policy, workplace culture and leadership capability may influence SLPs implementation of evidence-based practice recommendations (Cheung et al., 2013; Fulcher-Rood et al., 2018; Miao et al., 2015; Shrubsole et al., 2018; Young et al., 2018). One previous study identified that SLPs

employed in schools were more likely than SLPs in other settings to use evidence-based approaches when providing interventions for children (Koole, Nelson, & Curtis, 2015) and another identified that workplace policy appeared to influence the service delivery approaches chosen by SLPs more so than SLPs own theoretical or philosophical preferences (Zabiela, Leitão, & Williams, 2007). A previous study specifically investigating the language assessment practices used by SLPs employed in schools in the United States of America, found that workplace policy relating to eligibility for services was a key reason for use of norm-referenced assessments (Fulcher-Rood et al., 2018). However, as no previous surveys have specifically compared SLPs' child language assessments practices across different service agencies, further research is needed to understand the relationship between service agency and implementation of evidence-based practice recommendations for child language assessment.

*1.7.3.2. Child and family.* Previous studies have reported that parents' expectations of speech language pathology services may vary greatly (Donaldson, McDermott, Hollands, Copley, & Davidson, 2004). Some studies have identified that families value assessment services that are functional and relevant to a child's everyday performance (Donaldson et al., 2004; Roulstone, Wren, Bakopoulou, Goodlad, & Lindsay, 2015; Thomas-Stonell, Oddson, Robertson, & Rosenbaum, 2009), while other studies have identified a preference by parents for a medical model of service provision (Band et al., 2002; Carroll, 2010; Glogowska & Campbell, 2000). However, as no study has explored the extent to which SLPs identify child and family expectations as a challenge when conducting language assessment, further investigation is needed.

*1.7.3.3. Professional interactions.* Studies exploring the practice of SLPs in adult rehabilitation settings have identified that attitudes from peers and colleagues influence implementation of evidence-based practice recommendations (Miao et al., 2015; Shrubsole et

al., 2018). Similarly, it has been identified that health practitioners may be more successful in instigating positive changes to clinical practice when working as part of a group (Harding, 2014; Miao et al., 2015). To gain insight into the influence that professional environment may have on SLPs' language assessment practice more specifically, further information is needed regarding the main challenges SLPs in different agencies perceive in relation to child language assessment and the main sources of information that influence SLPs' knowledge of language assessment practice.

*1.7.3.4. Resources and incentives.* Previous studies have identified that limited time, limited material resources, and lack of incentives for evidence-based practices may influence SLPs' clinical practice (Cheung et al., 2013; Shrubsole et al., 2018; Young et al., 2018). For example, limited time to review literature was identified as a key factor that impacts on the use of evidence-based practices for children with Autism Spectrum Disorder (Cheung et al., 2013). Other studies have identified that use of ICTs for delivering services may be influenced by limited access to resources such as high-speed internet, limited technical support; and financial disincentives such as additional costs of purchasing computer equipment (Hill & Miller, 2012; Mashima & Doarn, 2009; Molini-Avejonas, Rondon-Melo, de La Higuera Amato, & Samelli, 2015). Challenges related to limited time and limited assessment materials have been reported by SLPs in previous surveys investigating child language assessment practice (Arias & Friberg, 2015; Beck, 1995; Guiberson & Atkins, 2012; Huang et al., 1997; Teoh et al., 2017; Westerveld & Claessen, 2014; Wilson et al., 1991). However, these surveys were focussed on specific groups of SLPs or specific clinical populations. Therefore, further investigation is needed to more specifically examine the degree to which this challenge is reported by a broad population of SLPs.

*1.7.3.5. Social, political and legal.* Literature has identified that ethical and legal issues may impact on clinical practice (Flottorp et al., 2013). For example, the use of ICTs for

delivering services may be influenced by barriers related to maintaining confidentiality, obtaining informed consent and ensuring high quality video and audio for accurate clinical judgements (Boisvert, Lang, Andrianopoulos, & Boscardin, 2010; Mashima & Doarn, 2009). Similarly, political issues such as funding allocations may influence clinical practice, for example, a lack of reimbursement from health insurers for services provided by ICTs has been reported as a barrier in literature (Mashima & Doarn, 2009). Collecting further information on the assessment practices SLPs use and the challenges SLPs experience in relation to child language assessment practice may help to identify social, political or legal issues that need to be considered to support future evidence-based assessment practice.

**1.8. Australian Speech Language Pathology Service Provision**

In Australia, no consistent guidelines for access to speech language pathology services exist at the federal level, resulting in differences in funding models, eligibility criteria and service provision across individual states/territories and service agencies (Speech Pathology Australia, 2011b). Service delivery is influenced by geographical location, for example, the states of Queensland, Victoria, Tasmania and South Australia provide SLP services through relevant public education departments; while the states of New South Wales, Western Australia and Northern Territory have very limited or no services provided through public education departments (Speech Pathology Australia, 2014). Children residing in these latter states access services through the public health system, however, access to services is variable and may become more limited as children reach school age (Speech Pathology Australia, 2014). Children in Australia may also access speech language pathology services through a range of other agencies, including publicly funded disability service providers, privately run clinics, private contractors employed by schools, non-government agencies and university teaching clinics. Large variations are recognised across states and service agencies

in terms service provision (Speech Pathology Australia, 2011b), however, limited information exists regarding the implications of these variations for clinical practice.

Furthermore, a new funding scheme called the National Disability Insurance Scheme (NDIS) is currently being rolled out across Australia (Australian Government, 2019). This scheme has provided a significant increase in funding for disability services and, consequently, has created a substantial increase in demand for allied health professionals including private SLP services (Commonwealth of Australia: Department of Social Services, 2019). This increase in service demand is expected to lead to growth in the development of new policy and practice standards for disability service provision (Commonwealth of Australia: Department of Social Services, 2019). The NDIS places particular emphasis on assessing service needs in relation to a person's ability to participate in meaningful activities of daily living (Speech Pathology Australia, 2019). Given the forthcoming growth in service demands, a focus on SLP assessment practice is needed to assist in identifying the future professional development needs of the Australian SLP workforce.

It is widely accepted that a wide range of cultural and linguistic diversity (CALD) exists across the Australian population (Williams & McLeod, 2012). The Australian Early Development Census (AEDC) collects data on children in Australia in their first year of full-time school (Australian Government, 2015). In recent findings from this census, the AEDC reported that 21.5% of children starting school in Australia speak a language other than Standard Australian English and a total of 331 different languages are spoken by children across Australia (Australian Early Development Census, 2015). Therefore, to gain an accurate profile of SLP practice in Australia, it is important that surveys of Australian SLP practice include the wide range of children with CALD backgrounds that are represented on typical SLP caseloads.

Only two previous surveys have specifically examined the language assessment practices Australian SLPs use for school-aged children (Westerveld & Claessen, 2014; Williams & McLeod, 2012). One survey investigated SLPs assessment practice specifically for children from CALD backgrounds (Williams & McLeod, 2012). This survey identified that although SLPs reported using a range of assessments, an overuse of English norm-referenced language measures appeared to exist for this population. The other survey specifically investigated assessments described as language sampling (Westerveld & Claessen, 2014). In this survey it was identified that most SLPs used language sampling, although data on the frequency with which they use this type of assessment was not reported on. Differences across states were also identified in this study, with SLPs in Queensland and Western Australia being more likely to report using language sampling measures with set guidelines for administration and scoring compared with SLPs in other states. Given the narrow scope of these surveys, further data are needed to understand the current landscape of Australian SLPs child language assessment practices more generally. Specifically, information examining differences across geographical locations and service agencies is needed to identify the influence of contextual factors on language assessment practice for school-aged children. This information will assist in identifying future directions for SLP assessment practice, including SLP professional development needs.

## 1.9. Terminology for Describing Language Assessments

The need for further surveys investigating SLPs assessment practice for school-aged children is apparent, however, lack of consistent terminology for describing clinical assessment practices has been identified as a barrier for accurate survey data collection (Pring, Flood, Dodd, & Joffe, 2012). This lack of consistent terminology is highlighted in a previous study examining the terms used by SLPs in case notes to describe assessments and interventions (Cowie et al., 2001). This study found that terms were not only used

inconsistently between different SLPs, but also between different case notes written by the same SLP. Furthermore, clinical practices in case notes were often not described in enough detail for others to easily understand or were described using terms that could be ambiguous in meaning.

Inconsistent use of terminology is also apparent in previous surveys of SLP assessment practice. Terms used in previous surveys to describe child language assessment practices are listed in Table 1.1. Some surveys included 'language sampling' and 'observations' as different response options, without appearing to give detailed explanations of the differences between these terms (Beck, 1995; Caesar & Kohler, 2007, 2009; Roulstone, Marshall, et al., 2015; Singh et al., 2016). Other surveys included 'standardised' and 'language sampling' as two different assessments, without clarifying how standardised narrative language sampling measures would be categorised within these two categories (Arias & Friberg, 2015; Caesar & Kohler, 2009; Fulcher-Rood et al., 2018; Pavelko et al., 2016; Singh et al., 2016). In one the survey by Roulstone, Marshall, et al. (2015), SLPs described their assessment practices using a variety of descriptors including: 'observation', 'play', 'audio-recording', 'videoing', 'language sampling in context', 'picture description', 'books' and 'posting games'. The use of such imprecise and ambiguous terminology makes it difficult to profile SLP's assessment practices with accuracy or detail.

To facilitate consistent and detailed descriptions from SLPs regarding the assessment practices they use, a taxonomy with well-defined and agreed-upon terminology is needed (Cowie et al., 2001). In addition to facilitating accurate data collection in survey research, such a taxonomy may enhance the effectiveness of professional communications and stimulate much-needed reflection on current practice, which is necessary for advancing clinical practice in the field of child language assessment (Eadie, 2003).

**1.10. Thesis Outline**

This thesis consists of five inter-related studies spanning three areas of research need. These studies collectively contribute to the overall thesis objective:

*To identify future actions and research directions that may facilitate implementation of evidence-based practice recommendations for child language assessment.*

Future directions for research may include a range of actions such as developing and providing professional development courses for SLPs, creating assessment policies or guidelines, creating resources and facilitating further data collection through research.

The first chapter of this thesis describes the background and structure of the thesis and outlines the objectives of the studies. The studies are then presented in Chapters 2-5. Detailed information is provided in these chapters regarding the existing literature, research methods, results and implications of each study. In the final chapter, knowledge gained from each study is synthesised and discussed within an implementation science framework. Limitations of the research and future directions are also discussed in the final chapter.

**1.10.1. Research area one: Psychometric properties of child language assessments.** The research conducted in research area one (reported in Chapter 2), addresses the need for studies examining the psychometric quality of comprehensive, norm-referenced spoken language measures for school-aged children. In this study, a systematic review was conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guideline (Liberati et al., 2009). The review included psychometric studies published in both assessment manuals and journal articles and the studies identified for inclusion in the review were rated for methodological quality using the COnsensus-based

Standards for the selection of health Measurement Instruments COSMIN checklist (Mokkink et al., 2010). Specifically, the objectives of the study in Chapter 2 were:

1. To determine the psychometric quality of currently available comprehensive norm-referenced spoken language measures for school-aged children (aged 4-12 years).

2. To identify the comprehensive norm-referenced spoken language measures with the best evidence for use.

**1.10.2. Research area two: Terminology for describing child language assessments.** The research conducted in research area two addresses the need for consensus on consistent and well-defined terminology for describing language assessment practice for school-aged children. Two studies were conducted in relation to research area two. In the first study (reported in Chapter 3), a taxonomy with terminology for describing language assessment practices was developed and presented to Australian SLPs with expertise in child language assessment during a three round Delphi consensus study. The specific objectives of this study were:

1. To develop a taxonomy that is agreed upon by experts and provides distinct, well-defined categories for describing language assessment practises for children.

2. To examine SLPs' application of a taxonomy for describing language assessment practices in clinical contexts.

Findings from the Delphi study identified a need for further investigation into strategies that may facilitate consistent application of the taxonomy by SLPs. In response to this, the second study (reported in Chapter 4) was conducted. Semi-structured interviews were used to gather perspectives from 13 Delphi study participants regarding the factors that may influence consistent application of the taxonomy. The specific objectives of this study were:

1. To identify SLPs' perceptions regarding factors that may influence consistent application of a taxonomy with terminology for describing language assessment practices for children (aged 4-18 years).

2. To identify SLPs' perceptions of strategies that may facilitate consistent use of the taxonomy when describing language assessment practices.

**1.10.3. Research area three: Language assessment practices used by Australian SLPs.** The research targeted at research area three addresses the need for data on current clinical child language assessment practice. A national survey of Australian SLPs was conducted. This survey consisted of two parts and findings are reported in two thesis chapters (Chapters 5 and 6). The terminology from the taxonomy developed in Chapters 3 and 4 was used to guide the survey questions and facilitate consistent interpretation of survey questions across participants.

Chapter 5 (Survey Part I) investigated the regularity with which SLPs use different types of assessments and the factors that influence regularity of use. This study also examined the main challenges SLPs experience in relation to child language assessment and the main sources from which they obtain information on language assessment practice. Specifically, the objectives were:

1. To identify the types of assessments SLPs use most regularly when assessing the language abilities of primary school-aged children (aged 4-12 years).

2. To identify if the following factors influence the regularity with which SLPs use different types of language assessments for primary school-aged children: service agency, years of working experience, SLPs' qualifications, proportion of children from CALD backgrounds on their caseload and geographical location in terms of Australian state and remoteness area classification.

3. To identify the challenges that SLPs in different agencies most frequently report in relation to child language assessment.

4. To identify the sources of information that SLPs from different agencies report most frequently for obtaining information on child language assessment.

Chapter 6 (Survey Part II) investigated the specific language measures, assessment procedures and assessment delivery methods used by SLPs when assessing the language abilities of school-aged children. The domains assessed, purposes of use and reasons for which language measures were chosen for use were also examined. Specifically, the objectives were:

1. To identify the specific language measures, assessment procedures and assessment delivery methods that SLPs use when assessing the language abilities of primary school children (aged 4-12 years) and the regularity with which these are used.

2. To identify the domains targeted, the purposes of use and the main reasons why regularly used language measures are chosen for use.

The objectives and research methods for each of the studies conducted in this thesis are also summarised in Figure 1.2.

**Area One Objectives:**

Examine the psychometric qualities of comprehensive norm-referenced spoken language measures for school-aged children

Identify assessments with the best psychometric evidence

**Area Two Objectives:**

Obtain consensus on terminology for describing language assessment practices for school-aged children

Identify strategies to facilitate consistent application of terminology by SLPs

**Area Three Objectives:**

Identify the types of assessments used, factors that influence use, challenges and sources of information

Identify the specific measures, procedures and delivery methods used and the domains, purposes and reasons for which language measures are used.

Method:

Systematic review and psychometric rating of language measures

Method:

Delphi Study and Semi-structured interviews

Method:

Online survey of Australian SLPs' assessment practices

Outcome: Psychometric ratings of language measures

Outcome: Agreed-upon terminology and strategies for application

Outcome: Profile of Australian SLPs' assessment practices

**Expected Outcome: Identify future actions and research directions that may facilitate implementation of evidence-based practice recommendations for child language assessment**

*Figure 1.2*. Overview objectives and research methods within the of thesis

**References for Chapter 1**

American Educational Research Association. (2014). Introduction. In *Standards for educational and psychological testing* (pp. 1-7). Washington, DC.

American Speech-Language-Hearing Association. (2000). Guidelines for the Roles and Responsibilities of the School-Based Speech-Language Pathologist [Guidelines]. Retrieved from www.asha.org/policy

American Speech and Hearing Association. (2018). *Assessment tools, techniques, and data sources* Retrieved from https://www.asha.org/Practice-Portal/Clinical-Topics/Late-Language-Emergence/Assessment-Tools-Techniques-and-Data-Sources/

Andersson, L. (2005). Determining the adequacy of tests of children's language. *Communication Disorders Quarterly, 26*(4), 207-225. doi:10.1177/15257401050260040301

Arias, G., & Friberg, J. (2015). Bilingual language assessment: Contemporary versus recommended practice in American schools. *Language, Speech, and Hearing Services in Schools, 48*(1), 1-15. doi:0.1044/2016_LSHSS-15-0090

Australian Government. (2015). Australian Early Development Census National Report: A Snapshot of Early Childhood Developnment in Australia. Retrieved from https://www.aedc.gov.au/resources/detail/2015-aedc-national-report

Band, S., Lindsay, G., Law, J., Soloff, N., Peacey, N., Gascoigne, M., & Radford, J. (2002). Are Health and Education talking to each other? Perceptions of parents of children with speech and language needs. *European Journal of Special Needs Education, 17*(3), 211-227. doi:10.1080/08856250210162121

Beck, A. R. (1995). Language assessment methods for three age groups of children. *Journal of Communication Development, 17*(2), 51-56. doi:10.1177/152574019501700206

Belkhodja, O., Amara, N., Landry, R., & Ouimet, M. (2007). The extent and organizational

determinants of research utilization in Canadian health services organizations. *Science*

*Communication, 28*(3), 377-417. doi:10.1177/1075547006298486

Betz, S. K., Eickhoff, J. K., & Sullivan, S. F. (2013). Factors influencing the selection of

standardized tests for the diagnosis of Specific Language Impairment. *Language,*

*Speech and Hearing Services in Schools, 44*, 133-143. doi:10.1044/0161-

1461(2012/12-0093)

Bishop, D. V. M., & McDonald, D. (2009). Identifying language impairment in children:

Combining language test scores with parental report. *International Journal of*

*Language & Communication Disorders, 44*(5), 600-615.

doi:10.1080/13682820802259662

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2016). CATALISE:

A multinational and multidisciplinary Delphi consensus study. Identifying language

impairments in children. *PLoS ONE, 11*(7), e0158753.

doi:10.1371/journal.pone.0158753

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE-2

consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary

Delphi consensus study of problems with language development: Terminology.

*Journal of Child Psychology and Psychiatry*. doi:10.1111/jcpp.12721

Boisvert, M., Lang, R., Andrianopoulos, M., & Boscardin, M. L. (2010). Telepractice in the

assessment and treatment of individuals with autism spectrum disorders: A systematic

review. *Developmental Neurorehabilitation, 13*(6), 423-432.

doi:10.3109/17518423.2010.499889

Bower, C., Watkins, R. E., Mutch, R. C., Marriott, R., Freeman, J., Kippin, N. R., . . . Giglia,

R. (2018). Fetal alcohol spectrum disorder and youth justice: A prevalence study

among young people sentenced to detention in Western Australia. *BMJ open, 8*(2), e019605. doi:10.1136/bmjopen-2017-019605

Bryan, K., Freer, J., & Furlong, C. (2007). Language and communication difficulties in juvenile offenders. *International Journal of Language Communication Disorders, 42*(5), 505-520. doi:10.1080/13682820601053977

Cabana, M. D., Rand, C. S., Powe, N. R., Wu, A. W., Wilson, M. H., Abboud, P. A. C., & Rubin, H. R. (1999). Why don't physicians follow clinical practice guidelines? A framework for improvement. *Journal of the American Medical Association, 282*(15), 1458-1465. doi:10.1001/jama.282.15.1458

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*, 190-200. doi:10.1044/0161-1461(2007/020)

Caesar, L. G., & Kohler, P. D. (2009). Tools clinicians use: A survey of language assessment procedures used by school-based speech-pathologists. *Communication Disorders Quarterly, 30*(4), 226-236. doi:10.1177/1525740108326334

Carroll, C. (2010). "It's not everyday that parents get a chance to talk like this": Exploring parents' perceptions and expectations of speech-language pathology services for children with intellectual disability. *International Journal of Speech-Language Pathology, 12*(4), 352-361. doi:10.3109/17549500903312107

Catts, H. W., Bridges, M. S., Little, T. D., & Tomblin, J. B. (2008). Reading achievement growth in children with language impairments. *Journal of Speech, Language, and Hearing Research, 51*, 1569-1579. doi:10.1044/1092-4388(2008/07-0259)

Cheung, G., Trembath, D., Arciuli, J., & Togher, L. (2013). The impact of workplace factors on evidence-based speech-language pathology practice for children with autism

spectrum disorders. *International Journal of Speech-Language Pathology, 15*(4), 396-406. doi:10.3109/17549507.2012.714797

Ciccia, A. H., Whitford, B., Krumm, M., & McNeal, K. (2011). Improving the access of young urban children to speech, language and hearing screening via telehealth. *Journal of telemedicine and telecare, 17*(5), 240-244. doi:10.1258/jtt.2011.100810

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*(8), 1106–1116. doi:10.1177/0956797610375449

Conti-Ramsden, G., Durkin, K., Simkin, Z., & Knox, E. (2009). Specific language impairment and school outcomes: Identifying and explaining variability at the end of compulsory education. *International Journal of Speech-Language Pathology, 44*(1), 15-35. doi:10.1080/13682820801921601

Conti-Ramsden, G., Simkin, Z., & Botting, N. (2006). The prevalence of autistic spectrum disorders in adolescents with a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 47*(6), 621-628. doi:10.1111/j.1469-7610.2005.01584.x

Cowan, R., Donlan, C., Newton, E. J., & Llyod, D. (2005). Number skills and knowledge in children with specific language impairment. *Journal of Educational Psychology, 97*(4), 732-744. doi:10.1037/0022-0663.97.4.732

Cowie, M., Wanger, K. M., Cartwright, A., Bailey, H., Millar, J. A., & Price, M. (2001). A review of Clinical Terms Version 3 (Read Codes) for speech and language record keeping. *International Journal of Language & Communication Disorders, 36*(1), 117-126. doi:10.1080/13682820116848

Crais, E. R., Roy, V. P., & Free, K. (2006). Parents' and professionals' perceptions of the implementation of family-centered practices in child assessments. *American Journal of Speech-Language Pathology, 15*(4), 365-377. doi:10.1044/1058-0360(2006/034)

Dethorne, L. S., Johnson, B. W., & Loeb, J. W. (2005). A closer look at MLU: What does it really measure? *Clinical Linguistics & Phonetics, 19*(8), 635-648. doi:10.1080/02699200410001716165

Dockrell, J. E., & Lindsay, G. (1998). The ways in which speech and language difficulties impact on children's access to the curriculum. *Child Language Teaching and Therapy, 14*(2), 117-133. doi:10.1177/026565909801400201

Dockrell, J. E., Lindsay, G., Letchford, B., & Mackie, C. (2006). Educational provision for children with specific speech and language difficulties: Perspectives of speech and language therapy service managers. *International Journal of Language & Communication Disorders, 41*(4), 423-440. doi:10.1080/13682820500442073

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. *Child and Adolescent Mental Health, 20*(2), 116-125. doi:10.1111/camh.12072

Donaldson, N., McDermott, A., Hollands, K., Copley, J., & Davidson, B. (2004). Clinical reporting by occupational therapists and speech pathologists: Therapists' intentions and parental satisfaction. *Advances in Speech Language Pathology, 6*(1), 23-38. doi:10.1080/14417040410001669471

Dopson, S., FitzGerald, L., Ferlie, E., Gabbay, J., & Locock, L. (2002). No magic targets! Changing clinical practice to become more evidence based. *Health care management review, 27*(3), 35-47. doi:10.1097/HMR.0b013e3181c88e79

Dunkley, C., Pattie, L., Wilson, L., & McAllister, L. (2010). A comparison of rural speech-language pathologists' and residents' access to and attitudes towards the use of

technology for speech-language pathology service delivery. *International Journal of Speech-Language Pathology, 12*(4), 333-343. doi:10.3109/17549500903456607

Durkin, K., Toseeb, U., Botting, N., Pickles, A., & Conti-Ramsden, G. (2017). Social confidence in early adulthood among young people with and without a history of language impairment. *Journal of Speech, Language, and Hearing Research, 60*(6), 1635-1647. doi:10.1044/2017_JSLHR-L-16-0256

Eadie, P. (2003). Speech pathology assessment practices: One assessment or many? *Advances in Speech Language Pathology, 5*(1), 65-68. doi:10.1080/14417040510001669081

Eadie, P., Conway, L., Hallenstein, B., Mensah, F., McKean, C., & Reilly, S. (2018). Quality of life in children with developmental language disorder. *International Journal of Language & Communication Disorders, 53*(4), 799-810. doi:10.1111/1460-6984.12385

Eadie, P., Nguyen, C., Carlin, J., Bavin, E., Bretherton, L., & Reilly, S. (2014). Stability of language performance at 4 and 5 years: measurement and participant variability. *International Journal of Language and Communication Disorders, 49*(2), 215-227. doi:10.1111/1460-6984.12065

Ebert, K. D., & Pham, G. (2017). Synthesizing information from language samples and standardized tests in school-age bilingual assessment. *Language, Speech, and Hearing Services in Schools, 48*(1), 42-55. doi:10.1044/2016_LSHSS-16-0007

Ebert, K. D., & Scott, C. M. (2014). Relationships between narrative language samples and norm-referenced test scores in language assessments of school-age children. *Language, Speech, and Hearing Services in Schools, 45*, 337–350. doi:10.1044/2014_LSHSS-14-0034

Edwards, M., Stredler-Brown, A., & Houston, K. T. (2012). Expanding use of telepractice in speech-language pathology and audiology *Volta Review, 112*(3), 227-242.

Feeney, R., Desha, L., Khan, A., & Ziviani, J. (2017). Contribution of speech and language difficulties to health-related quality-of-life in Australian children: A longitudinal analysis. *International Journal of Speech-Language Pathology, 19*(2), 139-152. doi:10.3109/17549507.2016.1151935

Finch, E., Cornwell, P., Ward, E. C., & McPhail, S. M. (2013). Factors influencing research engagement: Research interest, confidence and experience in an Australian speech-language pathology workforce. *BMC health services research, 13*(1), 144-155. doi:10.1186/1472-6963-13-144

Flottorp, S. A., Oxman, A. D., Krause, J., Musila, N. R., Wensing, M., Godycki-Cwirko, M., . . . Eccles, M. P. (2013). A checklist for identifying determinants of practice: A systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implementation Science, 8*(1), 1-11. doi:10.1186/1748-5908-8-35

Fraser, J., & Conti-Ramsden, G. (2008). Contribution of phonological and broader language skills to literacy. *International Journal of Language and Communication Disorders, 43*(5), 552-569. doi:10.1080/13682820701778069

Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact of diagnostic decisions? *Child Language Teaching and Therapy, 26*(1), 77-92. doi:10.1177/0265659009349972

Fulcher-Rood, K., Castilla-Earls, A. P., & Higginbotham, J. (2018). School-based speech-language pathologists' perspectives on diagnostic decision making. *American Journal of Speech-Language Pathology, 27*(2), 796-812. doi:10.1044/2018_AJSLP-16-0121

Gillon, G., Hyter, Y., Fernandes, F. D., Ferman, S., Hus, Y., Petinou, K., . . . Westerveld, M. (2017). International survey of speech-language pathologists' practices in working

with children with autism spectrum disorder. *Folia Phoniatrica et Logopaedica, 69*(1-2), 8-19. doi:10.1159/000479063

Glogowska, M., & Campbell, R. (2000). Investigating parental views of involvement in pre-school speech and language therapy. *International Journal of Language & Communication Disorders, 35*(3), 391-405. doi:10.1080/136828200410645

Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J., Caswell, W., & Robinson, N. (2006). Lost in knowledge translation: Time for a map? *The Journal of Continuing Education in the Health Professions, 26*, 13-24. doi:10.1002/chp.47

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

Hakkennes, S., & Dodd, K. (2008). Guideline implementation in allied health professions: A systematic review of the literature. *Quality and Safety in Healthcare, 17*, 296-300. doi:10.1136/qshc.2007.023804

Harding, K. E. (2014). Not enough time or a low priority? Barriers to evidence-based practice for allied health clinicians. *Journal of Continuing Education in the Health Professions, 34*(4), 224-231. doi:10.1002/chp.21255

Harlaar, N., DeThorne, L. S., Smith, J. M., Betancourt, M. A., & Petrill, S. A. (2016). Longitudinal effects on early adolescent language: A twin study. *Journal of Speech, Language, and Hearing Research, 59*(5), 1059-1073. doi:10.1044/2016_JSLHR-L-15-0257

Harrison, L. J., McLeod, S., Berthelsen, D., & Walker, S. (2009). Literacy, numeracy, and learning in school-aged children identified as having speech and language impairment in early childhood. *International Journal of Speech-Language Pathology, 11*(5), 392-403. doi:10.1080/17549500903093749

Hill, A. J., & Miller, L. E. (2012). A survey of the clinical use of telehealth in speech-language pathology across Australia. *Journal of Clinical Practice in Speech-Language Pathology, 14*(3), 110-117.

Hines, M., Lincoln, M., Ramsden, R., Martinovich, J., & Fairweather, C. (2015). Speech pathologists' perspectives on transitioning to telepractice: What factors promote acceptance? *Journal of telemedicine and telecare, 21*(8), 469-473. doi:10.1177/1357633X15604555

Hodge, M. A., Sutherland, R., Jeng, K., Bale, G., Batta, P., Cambridge, A., . . . Silove, N. (2018). Agreement between telehealth and face-to-face assessment of intellectual ability in children with specific learning disorder. *Journal of telemedicine and telecare*, Published early online. doi:10.1177/1357633X18776095

Hoffman, L. M., Ireland, M., Hall-Mills, S., & Flynn, P. (2013). Evidence-based speech-language pathology practices in schools: Findings from a national survey. *Language, Speech, and Hearing Services in Schools, 44*, 266-280. doi:10.1044/0161-1461(2013/12-0041)

Huang, R. J., Hopkins, J., & Nippold, M. A. (1997). Satisfaction with standardized language testing: A survey of speech-language pathologists. *Language, Speech, and Hearing Services in Schools, 28*(1), 12-29. doi:10.1044/0161-1461.2801.12

Hux, K., Morris-Friehe, M., & Sanger, D. D. (1993). Language Sampling Practices: A Survey of Nine States. *Language, Speech, and Hearing Services in Schools, 24*, 84-91. doi:10.1044/0161-1461.2402.84

Ireland, M., Hall-Mills, S., & Millikin, C. (2013). Appropriate implementation of severity ratings, regulations and state guidance: A responce to "Using norm-referenced tests to determine severity of language impairment in children: disconnect between U.S. policy makers and test developers" by Spaulding, Szulga, & Figueria. *Language,*

*Speech and Hearing Services in Schools, 44*, 320-323. doi:10.1044/0161-1461(2012/12-0048)

Johnson, C. J., Beitchman, J. H., & Brownlie, E. B. (2010). Twenty-year follow-up of children with and without speech-language impairments: Family, educational, occupational, and quality of life outcomes. *American Journal of Speech-Language Pathology, 19*, 51-65. doi:10.1044/1058-0360(2009/08-0083)

Kaderavek, J. N. (2015). Assessment of language disorders. In *Language disorders in children: Fundamental concepts of assessment and intervention* (pp. 45-95). US: Pearson.

Kelley, K., Clark, B., Brown, V., & Sitzia, J. (2003). Good practice in the conduct and reporting of survey research. *International Journal for Quality in health care, 15*(3), 261-266. doi:10.1093/intqhc/mzg031

Kemp, K., & Klee, T. (1997). Clinical language sampling practices: Results of a survey of speech-language pathologists in the United States. *Child Language Teaching and Therapy, 13*(2), 161-176. doi:10.1177/026565909701300204

Koole, H., Nelson, N. W., & Curtis, A. B. (2015). Factors influencing choices of contextualized versus traditional practices with children and adolescents who have traumatic brain injury. *Language, Speech, and Hearing Services in Schools, 46*, 352-361. doi:10.1044/2015_LSHSS-14-0109

Kritikos, E. P. (2003). Speech-language pathologists' beliefs about language assessment of bilingual/bicultural individuals. *American Journal of Speech-Language Pathology, 12*(1), 73-91. doi:10.1044/1058-0360(2003/054)

Lennox, M., Westerveld, M. F., & Trembath, D. (2018). Should we use sentence-or text-level tasks to measure oral language proficiency in year-one students following whole-class

intervention? *Folia Phoniatrica et Logopaedica, 69*(4), 169-179.

doi:10.1159/000485974

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., . . .
Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-
analyses of studies that evaluate health care interventions: Explanation and
elaboration. *PLoS medicine, 6*(7), e1000100. doi:10.1371/journal.pmed.1000100

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)

Lindsay, G., Dockrell, J., Letchford, B., & Mackie, C. (2002). Self esteem of children with
specific speech and language difficulties. *Child Language Teaching and Therapy,
18*(2), 125-143. doi:10.1191=0265659002ct231oa

Loucas, T., Charman, T., Pickles, A., Simonoff, E., Chandler, S., Meldrum, D., & Baird, G.
(2008). Autistic symptomatology and language ability in autism spectrum disorder
and specific language impairment. *Journal of Child Psychology and Psychiatry,
49*(11), 1184-1192. doi:10.1111/j.1469-7610.2008.01951.x

Lount, S. A., Purdy, S. C., & Hand, L. (2017). Hearing, auditory processing, and language
skills of male youth offenders and remandees in youth justice residences in New
Zealand. *Journal of Speech, Language, and Hearing Research, 60*(1), 121-135.
doi:10.1044/2016_JSLHR-L-15-0131

Mackie, C. J., Dockrell, J., & Lindsay, G. (2013). An evaluation of the written texts of
children with SLI: The contributions of oral language, reading and phonological
short-term memory. *Reading and Writing, 26*(6), 865-888. doi:10.1007/s11145-012-
9396-1

Maggio, V., Grañana, N. E., Richaudeau, A., Torres, S., Giannotti, A., & Suburo, A. M. (2014). Behavior problems in children with specific language impairment. *Journal of Child Neurology, 29*(2), 194-202. doi:10.1177/0883073813509886

Marshall, J., Goldbart, J., Pickstone, C., & Roulstone, S. (2015). Application of systematic reviews in speech-and-language therapy. *International Journal of Language & Communication Disorders, 46*(3), 261–272. doi:10.3109/13682822.2010.497530

Mashima, P. A., & Doarn, C. R. (2009). Overview of telehealth activities in speech-language pathology. *Telemedicine and e-Health, 14*(10), 1101-1117. doi:10.1089/tmj.2008.0080

McCauley, R. J., & Swisher, L. (1984). Psychometric reveiw of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders, 49*, 34-42. doi:10.1044/jshd.4901.34

McCormack, J., Harrison, L. J., McLeod, S., & McAllister, L. (2011). A nationally representative study of the association between communication impairment at 4–5 years and children's life activities at 7–9 years *Journal of Speech, Language, and Hearing Research, 54*, 1328-1348. doi:10.1044/1092-4388(2011/10-0155)

Miao, M., Power, E., & O'Halloran, R. (2015). Factors affecting speech pathologists' implementation of stroke management guidelines: A thematic analysis. *Disability and Rehabilitation, 37*(8), 674-685. doi:10.3109/09638288.2014.932444

Michie, S., Johnston, M., Abraham, C., Lawton, R., Parker, D., & Walker, A. (2005). Making psychological theory useful for implementing evidence based practice: A consensus approach. *Quality and Safety in Healthcare, 14*(1), 26-23. doi:10.1136/qshc.2004.011155

Mokkink, L. B., Prinsen, C. A. C., Bouter, L. M., De Vet, H. C. W., & Terwee, C. B. (2015). The COnsensus-based Standards for the selection of health Measurement INstruments

(COSMIN) and how to select an outcome measurement instrument. *Brazilian Journal of Physical Therapy, 20*(2), 105-113. doi:10.1590/bjpt-rbf.2014.0143

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., . . . de Vet, H. C. W. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology, 10*, 22-30. doi:10.1186/1471-2288-10-22

Molini-Avejonas, D. R., Rondon-Melo, S., de La Higuera Amato, C. A., & Samelli, A. G. (2015). A systematic review of the use of telehealth in speech, language and hearing sciences. *Journal of telemedicine and telecare, 21*(7), 367-376. doi:10.1177/1357633X15583215

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., . . . Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry, 57*(11), 1247-1257. doi:10.1111/jcpp.12573

O'Callaghan, A. M., McAllister, L., & Wilson, L. (2005). Barriers to accessing rural paediatric speech pathology services: Health care consumers' perspectives. *Australian Journal of Rural Health, 13*(3), 162-171. doi:10.1111/j.1440-1854.2005.00686.x

O'Connor, S., & Pettigrew, C. M. (2009). The barriers perceived to prevent the successful implementation of evidence-based practice by speech and language therapists. *International Journal of Language & Communication Disorders, 44*(6), 1018-1035. doi:10.3109/13682820802585967

Olswang, L. B., & Prelock, P. A. (2015). Bridging the gap between research and practice: Implementation science. *Journal of Speech, Language, and Hearing Research, 58*, 1818-1826. doi:10.1044/2015_JSLHR-L-14-0305

Paul, R., & Norbury, C. F. (2012). Chapter 2: Assessment. In R. Paul & C. F. Norbury (Eds.), *Language disorders from infancy through adolescence: Listening, speaking, reading, writing and communicating* (4th ed., pp. 22-60). St Louis, MI: Mosby Elsevier.

Pavelko, S. L., Owens, R. E., Ireland, M., & Hahs-Vaughn, D. L. (2016). Use of language sample analysis by school-based SLPs: Results of a nationwide survey. *Language, Speech, and Hearing Services in Schools, 47*(3), 246-258. doi:10.1044/2016_LSHSS-15-0044

Pearce, W. M., & Williams, C. (2013). The cultural appropriateness and diagnostic usefulness of standardized language assessments for Indigenous Australian children. *International Journal of Speech-Language Pathology, 15*(4), 429-440. doi:10.3109/17549507.2012.762043

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

Plante, E., & Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language, Speech, and Hearing Services in Schools, 25*(1), 15-24. doi:10.1044/0161-1461.2501.15

Pring, T., Flood, E., Dodd, B., & Joffe, V. (2012). The working practices and clinical experiences of paediatric speech and language therapists: A national UK survey. *International Journal of Language and Communication Disorders, 47*(6), 696-708. doi:10.1111/j.1460-6984.2012.00177.x

Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., . . . Wake, M. (2014). Terminological debate over language impairment in children: forward

movement and sticking points. *International Journal of Language & Communication Disorders, 49*(4), 452–462. doi:10.1111/1460-6984.12111

Renfrew, C. (2010). *Renfrew Action Picture Test*. Camberwell, Australia: ACER.

Roulstone, S. (2001). Consensus and variation between speech and language therapists in the assessment and selection of preschool children for intervention: A body of knowledge or idiosyncratic decisions? *International Journal of Language and Communication Disorders, 36*(3), 329–348. doi:10.1080/13682820010019928

Roulstone, S., Marshall, J., Powell, G. G., Goldbart, J., Wren, Y. E., Coad, J., . . . Coad, R. A. (2015). Evidence-based intervention for preschool children with primary speech and language impairments: Child Talk – an exploratory mixed-methods study. *Programme Grants for Applied Research, 3*(5), 1-408. doi:10.3310/pgfar03050

Roulstone, S., Wren, Y., Bakopoulou, I., Goodlad, S., & Lindsay, G. (2015). *Exploring interventions for children and young people with speech, language and communication needs: A study of practice*. Retrieved from https://www.gov.uk/government/publications/exploring-interventions-for-children-and-young-people-with-speech-language-and-communication-needs-a-study-of-practice

Ruggero, L., McCabe, P., Ballard, K. J., & Munro, N. (2012). Paediatric speech-language pathology service delivery: An exploratory survey of Australian parents. *International Journal of Speech-Language Pathology, 14*(4), 338-350. doi:10.3109/17549507.2012.650213

Santhanam, S. P., & Parveen, S. (2018). Serving culturally and linguistically diverse clients: A review of changing trends in speech-language pathologists' self-efficacy and implications for stakeholders. *Clinical Archives of Communication Disorders, 3*(3), 165-177. doi:10.21849/cacd.2018.00395

Shrubsole, K., Worrall, L., Power, E., & O'Connor, D. A. (2018). Barriers and facilitators to meeting aphasia guideline recommendations: What factors influence speech pathologists' practice? *Disability and Rehabilitation*, Advance online publication. doi:10.1080/09638288.2018.1432706

Siegel, E. B., Maddox, L. L., Ogletree, B. T., & Westling, D. L. (2010). Communication-based services for persons with severe disabilities in schools: A survey of speech-language pathologists. *Journal of Communication Disorders, 43*, 148-159. doi:10.1016/j.jcomdis.2009.12.003

Singh, J. S., Chan, M. Y., & Rusli, Y. A. (2016). Practise patterns of Malaysian speech-language pathologists in managing children with speech and language delay/disorder. *International Journal of Speech-Language Pathology, 18*(6), 560-570. doi:10.3109/17549507.2016.1139624

Snow, P. C., & Powell, M. B. (2008). Oral language competence, social skills and high-risk boys: What are juvenile offenders trying to tell us? *Children and Society, 22*, 16-28. doi:10.1111/j.1099-0860.2006.00076.x

Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Elligibility criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech and Hearing Services in Schools, 37*, 61-72. doi:10.1044/0161-1461(2006/007)

Spaulding, T. J., Szulga, M. S., & Figueroa, C. (2012). Using norm-referenced tests to determine severity of language impairment in children: Disconnect between US policy makers and test developers. *Language, Speech, and Hearing Services in Schools, 43*(2), 176-190. doi:10.1044/0161-1461(2011/10-0103)

Speech Pathology Australia. (2011a). Clinical guideline: Services in schools.

Speech Pathology Australia. (2011b). *Position paper: Speech pathology services in schools*. Retrieved from http://www.speechpathologyaustralia.org.au/

Speech Pathology Australia. (2014). *Submission to the inquiry into the prevalence of different types of speech, language and communication disorders and speech pathology services in Australia.* Retrieved from Australia: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/Speech_Pathology

Speech Pathology Australia. (2015). Clinical guideline: Evidence based speech pathology practice for individuals with Autism Spectrum Disorder.

Straus, S. E., Tetroe, J., & Graham, I. (2009). Defining knowledge translation. *Canadian Medical Association Journal, 181*(3-4), 165-168. doi:10.1503/cmaj.081229

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, 21*(1), 10-21. 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:10.3109/17549507.2011.65217

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & de Vet, H. C. W. (2012). Rating the methological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research, 21*, 651-657. doi:10.1007/s11136-011-9960-1

Thomas-Stonell, N., Oddson, B., Robertson, B., & Rosenbaum, P. (2009). Predicted and observed outcomes in preschool children following speech and language treatment:

Parent and clinician perspectives. *Journal of Communication Disorders, 42*(1), 29-42. doi:10.1016/j.jcomdis.2008.08.002

Thomas-Stonell, N., Washington, K., Oddson, B., Robertson, B., & Rosenbaum, P. (2013). Measuring communicative participation using the FOCUS©: Focus on the Outcomes of Communication Under Six. *Child: care, health and development, 39*(4), 474-480. doi:10.1111/cch.12049

Tucker, J. K. (2012). Perspectives of speech-language pathologists on the use of telepractice in schools: Quantitative survey results. *International Journal of Telerehabilitation, 4*(2), 61-72. doi:10.5195/ijt.2012.6100

Ukrainetz, T. A., & Blomquist, C. (2002). The criterion validity of four vocabulary tests compared with a language sample. *Child Language Teaching and Therapy, 18*(1), 59-78. doi:10.1191/0265659002ct227oa

Vallino-Napoli, L. D., & Reilly, S. (2004). Evidence-based health care: A survey of speech pathology practice. *Advances in Speech Language Pathology, 6*(2), 107-112. doi:10.1080/14417040410001708530

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010a). Assessment of children's literacy via an internet-based telehealth system. *Telemedicine and e-Health, 16*(5), 564-575. doi:10.1089/tmj.2009.0161

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010b). Internet-based telehealth assessment of language using the CELF–4. *Language Speech and Hearing Services in Schools, 41*, 445-458. doi:10.1044/0161-1461(2009/08-0131)

Watson, R. M., & Pennington, L. (2015). Assessment and management of the communication difficulties of children with cerebral palsy: A UK survey of SLT practice. *International Journal of Language & Communication Disorders, 50*(2), 241-259. doi:10.1111/1460-6984.12138

Westerveld, M. F., & Claessen, M. (2014). Clinician survey of language sampling practices in Australia. *International Journal of Speech-Language Pathology, 16*(3), 242-249. doi:10.3109/17549507.2013.871336

Westerveld, M. F., Gillon, G. T., & Boyd, L. (2012). Evaluating the clinical utility of the profile of oral narrative ability for 4-year-old children. *International Journal of Speech-Language Pathology, 14*(2), 130-140. doi:10.3109/17549507.2011.632025

Wiig, E. H., Semel, E., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals - 4th Edition* (5 ed.). Bloomington, MN: Pearson.

Williams, C., & McLeod, S. (2012). Speech-language pathologists assessment and intervention practices with multi-lingual children. *International Journal of Speech-Language Pathology, 14*(3), 292-305. doi:10.3109/17549507.2011.636071

Wilson, K. S., Blackmon, R. C., Hall, R. E., & Elcholtz, G. E. (1991). Methods of language assessment: A survey of California public school clinicians. *Language, Speech and Hearing Services in Schools, 22*, 236-241. doi:10.1044/0161-1461.2204.236

Yew, S. G. K., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry, 54*(5), 516-524. doi:10.1111/jcpp.12009

Yont, K. M., Hewitt, L. E., & Miccio, A. W. (2002). "What did you say?" Understanding conversational breakdowns in children with speech and language impairments. *Clinical Linguistics & Phonetics, 16*(4), 265-285. doi:10.1080/02699200210126523

Young, L., Shrubsole, K., Worrall, L., & Power, E. (2018). Factors that influence Australian speech-language pathologists' self-reported uptake of aphasia rehabilitation recommendations from clinical practice guidelines. *Aphasiology, 32*(6), 646-665. doi:10.1080/02687038.2018.1443201

Zabiela, C., Leitão, S., & Williams, C. (2007). Service delivery in rural, remote and regional

speech pathology. Part 2: The influence of context and philosophy. *ACQuiring*

*Knowledge in Speech, Language and Hearing, 9*(2), 44-47.

**Chapter 2.**

**Psychometric Properties of Language Assessments: A Systematic Review**

**Overview for Chapter 2 (Journal Article 1)**

Chapter 2 relates to research area one. This chapter details the results of a systematic review investigating the psychometric properties of comprehensive norm-referenced spoken language assessments for primary school-aged children. Norm-referenced, diagnostic language measures are frequently used to inform important decisions such as determining eligibility for services; therefore, these measures were selected as the focus of this review. Comprehensive language measures (i.e., measures that target multiple language domains) are suggested for use initially when assessing primary school-aged children, with more specific areas for assessment being identified from the initial assessment (Betz, Eickhoff, & Sullivan, 2013; Bishop, Snowling, Thompson, & Greenhalgh, 2016). Therefore, additional selection criteria included selection of measures that assess multiple language domains. As this review was conducted in 2014, only language measures published before this time are included.

This chapter 'sets the scene' by highlighting the strengths and weaknesses of available language measures. Findings from the review also provide information that informs evidence-based practice recommendations for use of diagnostic assessments for monolingual English-speaking children.

This chapter contains an Accepted Manuscript of an article published by Frontiers in Psychology on 07 September 2017, available online:

https://doi.org/10.3389/fpsyg.2017.01515. The spelling and wording contained within this chapter is that of the published manuscript. Note: In this manuscript the term 'language impairment' was used in place of 'language disorder' and the term 'standardized assessment' was used to refer to 'norm-referenced language measure'.

# Psychometric Properties of Language Assessments for Children Aged 4-12 Years:

# A Systematic Review

Deborah Denman[1*], Renée Speyer[1,2,3], Natalie Munro[4], Wendy M. Pearce[5], Yu-Wei Chen[4] ,

Reinie Cordier[2]

[1]College of Healthcare Sciences, James Cook University, Townsville, QLD, Australia

[2]Faculty of Health Science, Curtin University, Perth, WA, Australia

[3]Leiden University Medical Centre, Leiden, Netherlands

[4]Faculty of Health Science, The University of Sydney, Sydney, NSW, Australia

[5]School of Allied Health, Australian Catholic University, Sydney, NSW, Australia


*Corresponding author: E-mail: deborah.denman@postgrad.curtin.edu.au

**2.1. Abstract**

**Introduction:** Standardized assessments are widely used by speech pathologists in clinical and research settings to evaluate the language abilities of school-aged children and inform decisions about diagnosis, eligibility for services and intervention. Given the significance of these decisions, it is important that assessments have sound psychometric properties.

**Objective:** The aim of this systematic review was to examine the psychometric quality of currently available comprehensive language assessments for school-aged children and identify assessments with the best evidence for use.

**Methods:** Using the PRISMA framework as a guideline, a search of five databases and a review of websites and textbooks was undertaken to identify language assessments and published material on the reliability and validity of these assessments. The methodological quality of selected studies was evaluated using the COSMIN taxonomy and checklist.

**Results:** Fifteen assessments were evaluated. For most assessments evidence of hypothesis testing (convergent and discriminant validity) was identified; with a smaller number of assessments having some evidence of reliability and content validity. No assessments presented with evidence of structural validity, internal consistency or error measurement. Overall, all assessments were identified as having limitations with regards to evidence of psychometric quality.

**Conclusions:** Further research is required to provide good evidence of psychometric quality for currently available language assessments. Of the assessments evaluated, the Assessment of Literacy and Language, the Clinical Evaluation of Language Fundamentals- 5th Edition, the Clinical Evaluation of Language Fundamentals – Preschool: 2nd Edition and the Preschool Language Scales – 5th Edition presented with most evidence and are thus recommended for use.

**2.2. Introduction**

Language impairment refers to difficulties in the ability to comprehend or produce spoken language relative to age expectations (Paul & Norbury, 2012a). [1]Specific language impairment is defined when the language impairment is not explained by intellectual, developmental or sensory impairments (American Psychiatric Association, 2013; World Health Organisation, 2015). Specific Language Impairment is estimated to affect 2-10% of school-aged children with variation occurring due to using different diagnostic criteria (Dockrell & Lindsay, 1998; Law, Boyle, Harris, Harkness, & Nye, 2000; Lindsay, Dockrell, Desforges, Law, & Peacey, 2010). While there is active debate over terminology and definitions surrounding this condition (Ebbels, 2014), according to Bishop (Bishop, 2011), these children present with 'unexplained language problems' that require appropriate diagnosis and treatment because of their increased risk of long-term literacy difficulties (Catts, Bridges, Little, & Tomblin, 2008; Fraser & Conti-Ramsden, 2008), social-emotional difficulties (Conti-Ramsden & Botting, 2004; McCormack, Harrison, McLeod, & McAllister, 2011; Yew & O'Kearney, 2013)and poorer academic outcomes (Conti-Ramsden, Durkin, Simkin, & Knox, 2009; Dockrell & Lindsay, 1998; Harrison, McLeod, Berthelsen, & Walker, 2009)

Language assessments are used for a range of purposes. These include: initial screening, diagnosis of impairment, identifying focus areas for intervention, decision-making about service delivery, outcome measurement, epidemiological purposes and other research pursuits that investigate underlying cognitive skills or neurobiology (Paul & Norbury, 2012b; Shipley & McAfee, 2009; Tomblin, Records, & Zhang, 1996). Whilst few formal guidelines exist, current literature identifies that speech pathologists should use a range of assessment approaches when making judgements about the spoken language abilities of school-aged

children, such as: standardized assessment, language-sampling, evaluation of response-to-intervention, dynamic assessment, curriculum-based assessment and caregiver and teacher reports (Bishop & McDonald, 2009; Caesar & Kohler, 2009; Eadie et al., 2014; Friberg, 2010; Haynes & Pindzola, 2012; Hoffman, Leob, Brandel, & Gillam, 2011; Paul & Norbury, 2012c; Reed, 2005). Nonetheless, standardized assessments are a widely used component of the assessment process (Betz et al., 2013; Hoffman et al., 2011; Spaulding, Szulga, & Figueroa, 2012), particularly for determining if an individual meets diagnostic criteria for Language Impairment (American Psychiatric Association, 2013; World Health Organisation, 2015) and determining eligibility for services (Reed, 2005; Spaulding, Plante, & Farinella, 2006; Wiig, 2011). Standardized assessments are also designed to be easily reproducible and consistent, and as a result are also widely used in research (Betz et al., 2013; Tomblin et al., 1996)

Language assessments used in clinical practice and research applications must have evidence of sound psychometric properties (Andersson, 2005; Betz et al., 2013; Dockrell & Marshall, 2015; Terwee et al., 2012). Psychometric properties include the overarching concepts of validity, reliability and responsiveness (Mokkink, Terwee, Patrick, et al., 2010). This data is typically established by the developers of assessments and are often reported in the administration manuals for individual assessments (Hoffman et al., 2011). When data on psychometric properties is lacking, concerns may arise with the use of assessment results to inform important clinical decisions and the accuracy of reported outcome data in research (Friberg, 2010).

Previous studies have identified limitations with regards to the psychometric properties of spoken language assessments for school-aged children (Andersson, 2005; Friberg, 2010; McCauley & Swisher, 1984; Plante & Vance, 1994; Spaulding et al., 2006). An earlier study published in 1984 (McCauley & Swisher, 1984) examined the manuals of 30

speech and language assessments for children in relation to ten psychometric criteria. These criteria were selected by the authors and included description and size of normative sample, selection of items, normative data provided, concurrent and predictive validity, reliability and description of test administration. The appraisal indicated that only 20% of the 30 examined assessments met half of the criteria with most assessments meeting only two of the ten criteria. A decade later this information was updated through another study (Plante & Vance, 1994) examining the manuals of pre-school language assessments using the same ten criteria. In this later study, 38% of the 21 examined assessments met half the criteria with most assessments meeting four of the ten criteria.

More recently, literature has focussed on diagnostic accuracy (sensitivity and specificity). Although this information is often lacking in child language assessments, some authors have suggested that diagnostic accuracy should be a primary consideration in the selection of diagnostic language assessments, and have applied the rationale of examining diagnostic accuracy first when evaluating assessments (Friberg, 2010). A study published in 2006 (Spaulding et al., 2006) examined the diagnostic accuracy of 43 language assessments for school-aged children. The authors reported that 33 assessment manuals contained information to calculate mean differences between children with and without language impairment. While nine assessments included sensitivity and specificity data in the manual, only five of these assessments were determined by the authors as having an acceptable level of sensitivity and specificity (80% or higher). In another study published in 2010 (Friberg, 2010), an unspecified number of assessment manuals were examined with nine assessments identified as having an acceptable level of sensitivity and specificity. These nine assessments were then evaluated using 11 criteria based on a modification of the ten criteria used in earlier studies (McCauley & Swisher, 1984; Plante & Vance, 1994). No assessments were found to meet all 11 of the psychometric criteria, however all assessments met 8-10 criteria. The

findings from these studies suggest that, while the psychometric quality of assessments appears to have improved over the last 30 years, assessments of children's language may still require further development to improve their psychometric quality.

No previous reviews investigating the psychometric properties of language assessments for children were systematic in identifying assessments for review or included studies published outside of assessment manuals. This is important for two reasons, first, to ensure that all assessments are identified, and second, to ensure that all the available evidence for assessments, including evidence of psychometric properties that was published in peer reviewed journals, is considered when making overall judgements. Previous reviews have also lacked a method of evaluating the methodological quality of the studies selected for review. When evaluating psychometric properties, it is important to consider not only outcomes from studies, but also the methodological quality of studies. If the methodological quality of studies is not sound, then outcomes of studies cannot be viewed as providing psychometric evidence (Terwee et al., 2012). In addition, many of the assessments reviewed in previous studies have since been superseded by newer editions. Older editions are often not printed once new editions are released; therefore, an updated review is needed to examine the evidence for assessments that are currently available to speech-pathologists.

In the time since previous reviews of child language assessments were conducted, research has also advanced considerably in the field of psychometric evaluation (Mokkink, Prinsen, Bouter, De Vet, & Terwee, 2015; Polit, 2015). In 2010, the Consensus Based Standards for the Selection of Health Status Measurement Instruments (COSMIN) taxonomy (http://www.cosmin.nl) was developed through a Delphi study including fifty-seven international experts from disciplines including psychometrics, epidemiology and clinimetrics (Mokkink, Terwee, Knol, et al., 2010; Mokkink, Terwee, Patrick, et al., 2010). COSMIN aims to improve the selection of health-related measurement instruments by

clinicians and researchers through the provision of evidence-based tools for use when appraising studies examining psychometric quality (Mokkink et al., 2015). This includes provision of a checklist (http://www.cosmin.nl/COSMIN%20checklist.html) for rating the methodological quality of studies examining psychometric properties (Terwee et al., 2012). The COSMIN taxonomy and checklist has been utilised in a large number of systematic reviews (http://www.cosmin.nl/images/upload/files/Systematic%20reviews%20using%20COSMIN.pdf); however, has not yet been applied in the evaluation of the methodological quality of children's language assessments.

The COSMIN taxonomy describes nine measurement properties relating to domains of reliability, validity and responsiveness. Table 2.1 provides an overview and definition of all the COSMIN domains and measurement properties (Mokkink, Terwee, Patrick, et al., 2010). As the terminology in COSMIN is not always consistent with terms used throughout literature (Terwee, de Vet, & Mokkink, 2015), examples of terms that may be used across different studies is also given in Table 2.1.

**2.2.1. Study aim**. The aim of this study was to systematically examine and appraise the psychometric quality of diagnostic spoken language assessments for school-aged children using the COSMIN checklist (Mokkink, Terwee, Knol, et al., 2010; Mokkink, Terwee, Patrick, et al., 2010). Specifically, this study aimed to collect information on the overall psychometric quality of assessments and identify assessments with the best evidence of psychometric quality.

Table 2.1

*COSMIN Domains, Psychometric Properties, Aspects of Psychometric Properties and*

*Similar Terms based on Mokkink, Terwee, Patrick, et al. (2010).*

| Domain | Psychometric property (definition) | Examples of Terms used outside of COSMIN that may relate to measurement property |
|---|---|---|
| Reliability | Internal consistency (The degree of the interrelatedness between items) | Internal reliability Content sampling Conventional item analysis |
| | Reliability (Variance in measurements which is because of 'true' differences among clients) | Inter-rater reliability Inter-scorer reliability |
| | | Test-retest reliability Temporal stability Time sampling Parallel forms reliability |
| | Measurement error (Systematic and random error of a client's score that is not due to true changes in the construct to be measured) | Standard Error of Measurement |
| Validity | Content validity (The degree to which the content of an instrument is an adequate reflection of the construct to be measured) | n/a |
| | Construct validity (The degree to which scores are consistent with hypotheses based on the assumption that the instrument validly measures the construct to be measured) | n/a |
| | Aspect of construct validity –structural validity (The degree to which scores reflect the dimensionality of the measured construct) | Internal structure |
| | Aspect of construct validity – hypothesis testing (Item construct validity) | Concurrent validity Convergent validity Predictive validity |
| | | Discriminant validity Contrasted groups validity Identification accuracy Diagnostic accuracy |
| | Aspect of construct validity- Cross cultural validity (The degree to which the performance of the items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version of the instrument) | n/a |
| | Criterion validity (The degree to which scores reflect measurement from a 'gold standard') | Sensitivity/specificity (when comparing assessment with gold-standard) |
| Responsiveness | Responsiveness (The ability to detect change over time in the construct to be measured) | Sensitivity/specificity (when comparing two administrations of an assessment) Changes over time Stability of diagnosis |
| [a] Interpretability | Interpretability (The degree to which qualitative meaning can be assigned to quantitative scores obtained from the assessment) | n/a |

Notes: [a] Interpretability is not considered a psychometric property

**2.3. Methods**

    **2.3.1. Selection criteria**. Assessments selected for inclusion in the review were standardized norm-referenced spoken language assessments from any English-speaking country with normative data for use with mono-lingual English-speaking children aged 4-12 years. Only the most recent editions of assessments were included. Initial search results indicated 76 assessments meeting this criterion. As it was not possible to review such a large number of assessments, further exclusion criteria were applied. Assessments were excluded if they were not published within the last 20 years. It is recognised that norm-referenced assessments should only be used with children whose demographics are represented within the normative sample (Friberg, 2010; Hegde & Pomaville, 2013; Paul & Norbury, 2012b); therefore the use of assessments normed on populations from several decades ago may be questionable with current populations. Screening assessments were excluded as they are designed to identify individuals who are at risk or may require further diagnostic assessment (Paul & Norbury, 2012b; Reed, 2005) and thus have a different purpose to diagnostic assessments. Similarly, assessments of academic achievement were also excluded, as although they may assess language ability, this occurs as part of a broad purpose of assessing literacy skills for academic success (Wiig, 2011).

    For diagnosis of Specific Language Impairment using standardized testing, previous research has recommended the use of composite scores that include measures of both comprehension and production of spoken language across three domains: word (semantics), sentence (morphology and syntax) and text (discourse) (Gillam, Peña, Bedore, Bohman, & Mendez-Perezb, 2013; Tomblin et al., 1996). While phonology and pragmatics may also be assessed, these areas are not typically considered part of the diagnostic criteria for identifying Specific Language Impairment (Tomblin et al., 1996). While some evidence suggests that children's language skills may not be contrastive across modalities of comprehension and

production (Leonard, 2009; Tomblin & Zhang, 2006), current literature conceptualises language in this way (Wiig, 2011; World Health Organisation, 2015). A recent survey of SLP's in the United States also identified that 'comprehensive' language assessments that assess multiple language areas are used more frequently than assessments that assess a single domain or modality (Betz et al., 2013). As comprehensive assessments provide a broad picture of a child's language strengths and weaknesses, these assessments are often selected first, with further examination of specific domains or modalities conducted if necessary (Betz et al., 2013; Dockrell & Marshall, 2015).

Given the support in literature for the use of comprehensive assessments in diagnostics and the wide use of these assessments by speech pathologists, it was identified that a review of comprehensive language assessments for school-aged children is of particular clinical importance. Therefore, assessments were included in this study if they were the latest edition of a language assessment with normative data for monolingual English speaking children aged 4-12 years; were published within the last 20 years; were primarily designed as a diagnostic assessment; and were designed to assess language skills across at least two of the following three domains of spoken language: word (semantics), sentence (syntax/morphology) and text (discourse).

**2.3.2. Sources of information**. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were developed through consensus of an international group to support high quality reporting of the methodology of systematic reviews (Moher, Liberati, Tetzlaff, & Altman, 2009) and were thus used to guide this review. Language assessments were identified through database searches and through comprehensively searching publisher websites, speech pathology websites and textbooks. A flowchart outlining sources of information is contained in Figure 2.1.
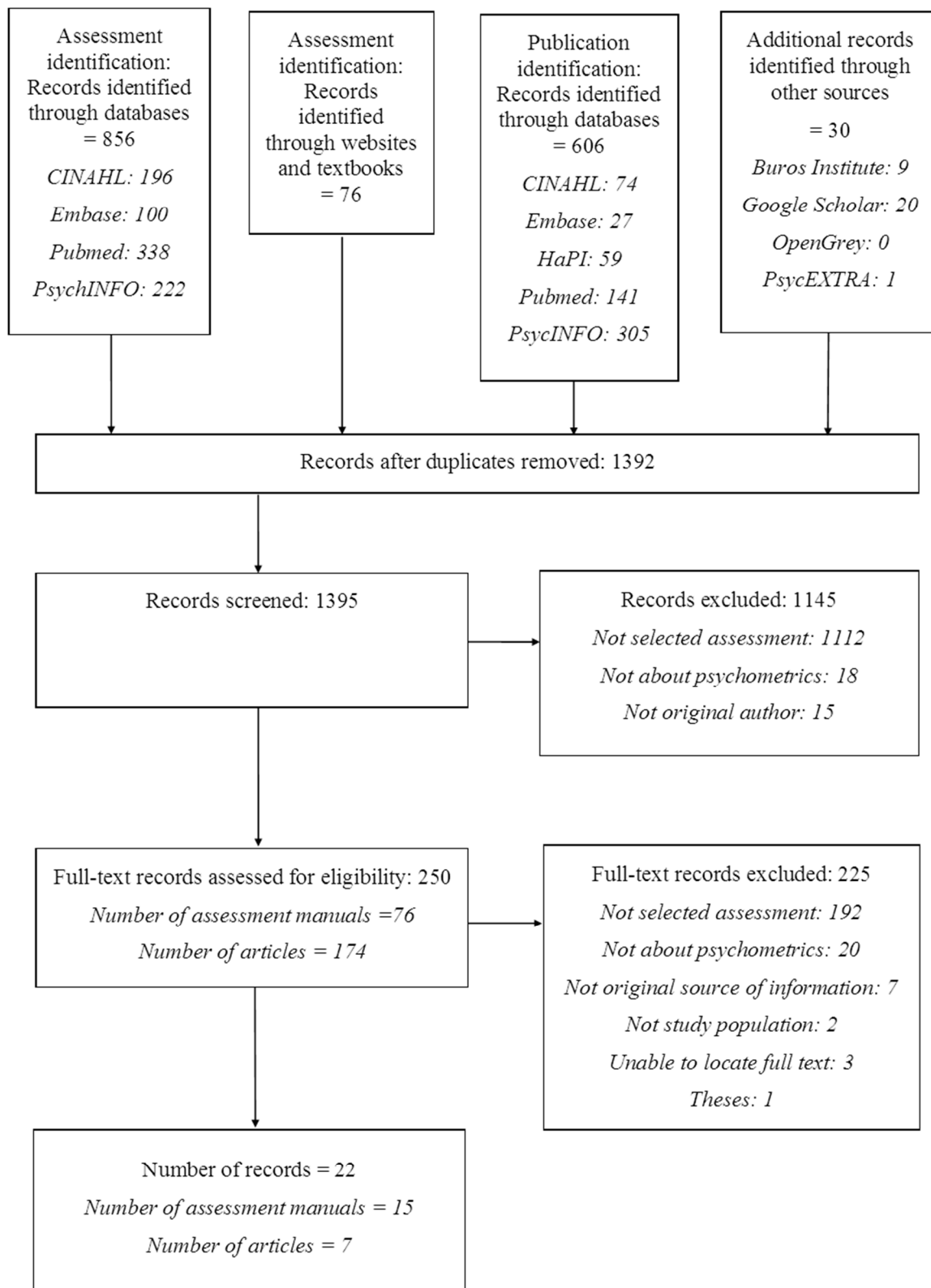
*Figure 2.1.* Flowchart of selection process according to PRISMA

Database searches of PubMed, CINAHL, PsycINFO and Embase were conducted between February and March 2014. Database searches were conducted with subject headings or mesh terms to identify relevant articles up until the search date. Free text word searches were also conducted for the last year up until the search date to identify recently published articles not categorised in subject headings. The search strategies are described in Table 2.2.

Assessments were also identified from searches of websites and textbooks. Speech pathology association websites from English speaking countries were searched and one website, the American Speech and Hearing Association, was identified as having an online directory of assessments. The website for this directory was identified as being no longer available as of 30/01/16. Publisher websites were identified by conducting Google searches with search terms related to language assessment and publishing and by searching the publisher sites from assessments already identified. These search terms are listed in Table 2.2. From these methods, a total of 43 publisher websites were identified and searched. Textbooks were identified from Google searches related to language assessment and the contents of recently published books searched. Three recently published textbooks (Hegde & Pomaville, 2013; Kaderavek, 2011; Paul & Norbury, 2012b) were identified as having lists of language assessments, which were then searched for assessments not already identified.

Published articles relating to psychometric properties of selected assessments were identified through additional database searches conducted between December 2014 and January 2015 using PubMed, CINAHL, Embase, PsycINFO and HaPI. Searches were conducted using full names of assessments as well as acronyms; and limited to articles written in English and published in or after the year the assessment was published. Articles were included in the psychometric evaluation if they related to one of the selected assessments, contained information on reliability and validity and included children speaking English as a first language in the study. Google Scholar, OpenGrey (http://www.opengrey.eu)

and PsycEXTRA® (http://www.apa.org/pubs/databases/psycextra/) were also searched for grey literature. Search terms are contained in Table 2.2.

All retrieved articles were reviewed for inclusion by two reviewers independently using selection criteria, with differences in opinion settled by group discussion to reach consensus. All appropriate articles up until the search dates were included.

**2.3.3. Study selection**. Across all searches, a total of 1,395 records were retrieved from databases and other sources. The abstracts for these records were reviewed and 1,145 records were excluded as they were not related to language assessment for mono-lingual English-speaking children aged 4-12 years. The full text versions of remaining records were then reviewed and 225 records were excluded as they did not provide information on the 15 selected assessments, did not contain information on the reliability and validity of selected assessments, did not examine the study population, or were unpublished or unable to be located. Records were also excluded if they were not an original source of information on the reliability and validity of selected assessments. For example, articles reviewing results from an earlier study or reviewing information from an assessment manual were not included if they did not contain new data from earlier studies. A total of 22 records were identified for inclusion, including 15 assessment manuals and 7 articles. Figure 2.1 represents the assessment and article selection process using a PRISMA flowchart.

Table 2.2

*Search Terms used in Database Searches*

| Assessment Identification | Database (Search Date) and Search Terms | Limitations |
|---|---|---|
| **Subject Headings** | **CINAHL (17.02.14):** ((MH 'Psychometrics') OR (MH "Measurement Issues and Assessments") OR (MH 'Reliability & Validity")) AND ((MH "Language tests") OR (MH "Speech and Language Assessment")) | Child, preschool: 2-5 years; Child: 6-12 years |
| | **Embase (17.02.14):** (psychometry/ OR validity/ OR reliability/) AND (Language test/) | English language; Preschool child <1 to 6 years>; School child <7 to 12 years> |
| | **PsycINFO (17.02.14):** ((DE "Psychometrics") OR (DE "Statistical reliability") OR (DE "Test reliability") OR (DE "Statistical validity") OR (DE "Test validity")) AND (DE "Language") AND ((DE "Testing") OR (DE "Measurement")) | No limitations |
| | **PubMed (17.02.14):** ("Psychometrics"[Mesh] OR "Reproducibility of Results"[Mesh]) OR "Validation Studies"[Publication Type] OR "Validation Studies as Topic"[Mesh]) AND ("Language Tests"[Mesh]) | (English[lang]) AND ('child'[MeSH Terms:noexp] OR 'child, preschool'[MeSH Terms]) |
| **Free Text Words** | **CINAHL (24.03.14):** (Psychometric* OR Reliability OR Validity) AND (Language OR Speech OR Vocabulary OR Grammar) AND (Measurement* OR Test OR Tests OR Testing OR Assessment* OR Screening*) | English language; Child, preschool: 2-5 years; Child: 6-12 years; Publication date: 20130101-20141231 |
| | **Embase (24.03.14):** *As per CINAHL Free Text* | English language; Preschool child <1 to 6 years>; School child <7 to 12 years>; yr='2013-Current' |
| | **PsycINFO (24.03.14):** *As per CINAHL Free Text* | English; Preschool age (2-5 years); School Age (6-12 years); Adolescence (13-17 years); Publication year: 2013-2014 |
| | **PubMed (17.02.14):** *As per CINAHL Free Text* | English; Preschool Child: 2 – 5 years; Child: 6 – 12 years; Publication date from 2013/01/01 to 2014/02/31 |
| **Grey Literature** | **Google (20:06:15):** ("Speech Pathology" OR "Speech Therapy" OR "Speech Language" AND "Assessment" OR "Test" AND "Publishers" OR "Publishing Companies" OR "textbooks" | No limitations |
| | **Speechbite (20/06/15):** "Language" AND "Assessment" OR "Test" OR "Psychometrics | No limitations |

| Publication Indentifications | Database (Search Date) and Search Terms | Limitations |
|---|---|---|
| **Free Text Words** [a] | **CINAHL (20.01.15):** (Name of assessment) OR (Acronym of assessment) | English Language |
| | **Embase (12.12.14):** *As per CINAHL Free Text* | English language |
| | **PsycINFO (22.01.15):** *As per CINAHL Free Text* | English |
| | **PubMed (23.01.15):** *As per CINAHL Free Text* | English |
| | **HaPI (06.12.14):** *As per CINAHL Free Text* | English |
| | **HaPI (06.12.14):** *As per CINAHL Free Text* | English |
| | **PsycEXTRA (21/01/15):** (Name of assessment) OR (Acronym of assessment) | Publication year of assessment to current |
| **Grey literature**[a] | **OpenGrey (22/01/15):** (Name of assessment) OR (Acronym of assessment) | No limitations |
| | **Google Scholar (11/01/15):** (Name of assessment) OR (Acronym of assessment) | No limitations |

Notes: [a] The title of the assessment and its acronym were used as search strategy.

**2.3.4. Data collection process and data synthesis**. Studies selected for inclusion in

the review were rated on methodological quality using COSMIN with the outcome from

studies then rated against criteria based on Terwee et al. (2007) and Schellingerhout,

Heymans, Verhagen, de Vet, and Terwee (2011). Studies for each measurement property for

each assessment were then combined to give an overall evidence rating for each assessment

using criteria based on Schellingerhout et al. (2011). This methodology is similar to

methodology used in previous systematic reviews examining the other health related

measurement instruments (Schellingerhout et al., 2011; Uijen et al., 2012; Vrijman et al.,

2012).

The four point COSMIN checklist

(http://www.cosmin.nl/images/upload/files/COSMIN%20checklist%20with%204-

point%20scale%2022%20juni%202011.pdf) was used for rating methodology (Terwee et al.,

2012). This checklist provides a system for rating each of the nine COSMIN measurement

properties (internal consistency, reliability, measurement error, content validity, structural

validity, hypothesis testing, cross-cultural validity, criterion validity and responsiveness).

Interpretability can also be measured but is not considered a psychometric property (Mokkink

et al., 2009). Each COSMIN measurement property is assessed on 5–18 items that rate the

standard of methodological quality using an 'excellent', 'good', 'fair' or 'poor' rating scale

(Terwee et al., 2012). Items vary depending on the property being rated; however, most

properties include ratings for reporting and handling of missing information, sample size,

design flaws and type of statistical analysis. There are also property specific items; for

example, time interval, patient stability and similarities in testing conditions are rated for test-

retest reliability studies.

Different methods for scoring the COSMIN 4-point checklist are employed in studies

examining the methodology of psychometric studies. One suggested method is a 'worst rating

counts' system, where each measurement property is given the score of the item with the

lowest rating (Terwee et al., 2012). The advantage of this method over other methods, such as

giving a 'mean score' for each measurement property, is that serious flaws cannot be compensated for by higher scores on other items (Terwee et al., 2012). However, the 'worst rating counts' system is severe as an assessment needs only one 'poor' rating to be 'poor' for a given measurement property and must receive all 'excellent' scores to be rated 'excellent' for a measurement property. Previous studies (Speyer, Cordier, Kertscher, & Heijnen, 2014) have also identified that this method lacks the ability to distinguish 'better' assessments when all reviewed assessments have limitations leading to poor ratings on some items.

In this current study, the scores for each item were averaged to give an overall rating for each measurement property. This provides information on the methodological quality in general for studies that were rated. In the scoring process, the appropriate measurement properties were identified and rated on the relevant items. The options for 'excellent', 'good', 'fair' and 'poor' on the 4-point checklist were ranked numerically, with 'excellent' being the highest score and 'poor' being the lowest score. As the current version of the COSMIN 4-point scale was designed for a 'worst rating counts' method, some items do not have options for 'fair' or 'poor'. Therefore, this was adjusted for in the percentage calculation so that the lowest possible option for each item was considered a 0 score. As each measurement property has a different number of items or may have items that are not applicable to a particular study, the number of items rated may differ across measurement properties or across studies. Therefore, overall scores for each measurement property rated from each study were calculated as a percentage of points received compared to total possible points that a study could have received for that measurement property. The resulting percentages for each measurement property were then classified according to quartile, that is: 'Poor' = 0-25%, 'Fair' = 25.1-50%, 'Good' = 50.1-75% and 'Excellent' = 75.1-100% (Cordier et al., 2015). Where a measurement property was rated 'excellent' or 'good' overall but had a 'poor' score at item level for important aspects such as sample size or statistical analysis, this was noted

so that both quantitative scores depicting overall quality and descriptive information about specific methodological concerns may be considered when viewing results.

The findings from studies with 'fair' or higher COSMIN ratings were subsequently appraised using criteria based on Terwee et al. (2007) and Schellingerhout et al. (2011). These criteria are described in Table 2.3. Because the COSMIN ratings were averaged to give a rating of overall quality and Table 2.3 rates studies against specific methodological criteria, it is possible for studies with good COSMIN ratings in to be rated as indeterminate from Table 2.3.

Overall evidence ratings for each measurement property for each assessment were then determined by considering available evidence from all the studies. These ratings were assigned based on the quality of the methodology of available studies (as rated using COSMIN) and the quality of the findings from the studies (as defined in Table 2.3). This rating scale was based on criteria used by Schellingerhout et al. (2011) and is outlined in Table 2.4**.**

Table 2.3

*Criteria for Measuring Quality of Findings for Studies Examining Measurement Properties based on Terwee et al. (2007) and Schellingerhout et al. (2011)*

| COSMIN measurement property | Rating | Quality Criteria |
|---|---|---|
| **Internal Consistency** | + | Subtests one-dimensional (determined through factor analysis with adequate sample size) and Cronbach alpha between 0.70-0.95 |
| | ? | Dimensionality of subtests unknown (no factor analysis) or Cronbach's alpha not calculated |
| | – | Subtests uni-dimensional (determined through factor analysis with adequate sample size) and Cronbach's alpha < 0.7 or >0.95 |
| | ± | Conflicting results |
| | NR | No information found on internal consistency |
| | NE | Not evaluated due to 'poor' methodology rating on COSMIN |
| **Reliability** | + | ICC/weighted Kappa equal to or > than 0.70 |
| | ? | Neither ICC/weighted Kappa calculated or doubtful design or method (e.g. time interval not appropriate) |
| | - | ICC/weighted Kappa < 0.70 with adequate methodology |
| | ± | Conflicting results |
| | NR | No information found on reliability |
| | NE | Not evaluated due to 'poor' methodology on COSMIN |
| **Measurement Error** | + | MIC > SDC or MIC equals or inside LOA |
| | ? | MIC not defined or doubtful design or method |
| | - | MIC < SDC or MIC equals or inside LOA with adequate methodology |
| | + | Conflicting results |
| | NR | No information found on measurement error |
| | NE | Not evaluated due to 'poor' methodology on COSMIN |
| **Content Validity** | + | Good methodology (i.e., an overall rating of 'Good' or above on COSMIN criteria for content validity) and experts examined all items for content and cultural bias during development of assessment |
| | ? | Questionable methodology or experts only employed to examine one aspect (e.g., cultural bias) |
| | - | No expert reviewer involvement |
| | ± | Conflicting results |
| | NR | No information found on content validity |
| | NE | Not evaluated due to 'poor' methodology |
| **Structural validity** | + | Factor analysis performed with adequate sample size. Factors explain at least 50% of variance |
| | ? | No factor analysis or inadequate sample size. Explained variance not mentioned |
| | - | Factors explain < 50% of variance despite adequate methodology |
| | ± | Conflicting results |
| | NR | No information found on structural validity |
| | NE | Not evaluated due to 'poor' methodology |
| **Hypothesis testing** | + | Convergent validity: Correlation with assessments measuring similar constructs equal to or >0.5 and correlation is consistent with hypothesis. Discriminant validity: findings consistent with hypotheses using appropriate statistical analysis (e.g., *t*-test p < 0.05 or Cohen's d effect size > 0.5) |
| | ? | Questionable methodology e.g. only correlated with assessments that are not deemed similar |
| | - | Discriminant validity: findings inconsistent with hypotheses (e.g., no significant difference identified from appropriate statistical analysis). Convergent validity: Correlation with assessments measuring similar constructs equal to or <0.5 or correlation is inconsistent with hypothesis |
| | ± | Conflicting results |
| | NR | No information found on hypothesis testing |
| | NE | Not evaluated due to 'poor' methodology |

Notes: + Positive result; - Negative result; ? Indeterminate result due to methodological shortcomings; -/+ Conflicting results within the same study (e.g., high correlations for some results but not on others); NR Not reported; NE Not evaluated; MIC minimal important change; SDC smallest detectable change; LOA limits of agreement; ICC Intra-class correlation; SD standard deviation.

Table 2.4

*Level of Evidence for Psychometric Quality for each Measurement Property based on*

*Schellingerhout et al. (2011)*

| Level of evidence | Rating | Criteria based on appraisal of quality of methodology (rated according to COSMIN) and quality of findings (rated according to Table 2.3) |
|---|---|---|
| **Strong evidence** | +++or --- | Consistent findings across 2 or more studies of 'good' methodological quality OR one study of 'excellent' methodological quality |
| **Moderate evidence** | ++ or -- | Consistent findings across 2 or more studies of 'fair' methodological quality OR one study of 'good' methodological quality |
| **Weak evidence** | + or - | One study of 'fair' methodological quality (examining convergent or discriminant validity if rating hypothesis testing) |
| **Conflicting evidence** | ± | Conflicting findings across different studies (i.e., different studies with positive and negative findings) |
| **Unknown** | ? | Only available studies are of 'poor' methodological quality |
| **Not Evaluated** | NE | Only available studies are of 'poor' methodological quality as rated on COSMIN |

Notes: + Positive result; - Negative result

To limit the size of this review, selected assessments were not appraised on the measurement property of responsiveness, as that would have significantly increased the size of the review. Interpretability is not considered a psychometric property and was also not reviewed. However, given the clinical importance of responsiveness and interpretability, it is recommended that these properties be a target for future research. Cross-cultural validity applies when an assessment has been translated or adapted from another language. As all the assessments reviewed in this study were originally published in English, cross-cultural validity was not rated. However, it is acknowledged that the use of English language assessments with the different dialects and cultural groups that exist across the broad range of English-speaking countries is an area that requires future investigation. Criterion validity was also not evaluated in this study as this measurement property refers to a comparison of an assessment to a diagnostic 'gold-standard' (Mokkink et al., 2010a). Consultation with experts and reference to current literature (Betz et al., 2013; Dollaghan & Horner, 2011; Tomblin et

al., 1996) did not identify a 'gold-standard' or an industry recognised 'reference standard' for diagnosis of language impairment, therefore all studies comparing one assessment to another assessment were considered convergent validity and rated as hypothesis testing according to COSMIN.

Diagnostic accuracy, which includes sensitivity and specificity and positive predictive power calculations, is an area that does not clearly fall into a COSMIN measurement property. However, current literature identifies this as being an important consideration for child language assessment (Friberg, 2010; Spaulding et al., 2006). In this review, data from studies examining diagnostic accuracy was collated in Table 2.9 to allow for this information be considered alongside information on COSMIN measurement properties. It should be noted that these studies were not rated for methodological quality, as the COSMIN checklist was not identified as providing an appropriate rating scale for these types of studies. However, descriptive information on the methodological quality of these studies was commented upon in the results section.

Where several studies examining one measurement property were included in a manual, one rating was provided based on information from the study with the best methodology. For example, if a manual included internal consistency studies using different populations then a rating for internal consistency was given based on the study with the most comprehensive or largest sample size. The exception was for reliability, where test-retest and inter-rater reliability were rated separately and hypothesis testing where convergent validity and discriminant validity were rated separately. In most cases, these different reliability and hypothesis testing studies were conducted using different sample sizes and different statistical analyses. As it was considered that manuals that include both these studies for each measurement property are providing evidence across different aspects of the measurement

property, it was decided that counting these as different studies would allow this to be reflected in final data.

Some assessments also included studies for hypothesis testing examining gender, age and socio-cultural differences. Whilst this information contributes important information on an assessment's usefulness, we identified convergent validity and discriminant validity as key aspects for the measurement property of hypothesis testing and thus only included these studies in this review.

**2.3.5. Risk of bias.** All possible items for each assessment were rated from all identified publications. Where an examination of a particular measurement property was not reported in a publication or not reported with enough detail to be rated, this was rated as 'not reported' (NR). Two raters were involved in appraising publications. To ensure consistency, both raters involved in the study trained as part of a group prior to rating the publications for this study. The first rater rated all publications with a random sample of 40% of publications also rated independently by a second rater. Inter-rater reliability between the two raters was calculated and determined to be adequate (weighted Kappa = 0.891; SEM=0.020; 95% confidence interval = 0.851-0.931). Any differences in opinion were discussed and the first rater then appraised the remaining 60% of articles applying rating judgements agreed upon after consensus discussions.

## 2.4. Results

**2.4.1. Assessments selected for review**. A total of 22 publications were identified for inclusion in this review. These included 15 assessment manuals and seven journal articles relating to a total of 15 different assessments. From the 22 publications, 129 eligible studies were identified, including three studies that provided information on more than one of the 15 selected assessments. Eight of these 129 studies reported on diagnostic accuracy and were included in the review, but were not rated using COSMIN, leaving 121 articles to be rated for

methodological quality. Of the 15 selected assessments, six were designed for children younger than 8 years and included the following assessments: Assessment of Literacy and Language (ALL; nine studies), Clinical Evaluation of Language Fundamentals: Preschool – 2nd Edition (CELF:P-2; 14 studies), Reynell Developmental Language Scales – 4th Edition (NRDLS; six studies), Preschool Language Scales – 5th Edition (PLS-5; nine studies), Test of Early Language Development – 3rd Edition (TELD-3; nine studies) and Test of Language Development – Primary: 4th Edition (TOLD-P:4; nine studies). The Test of Language Development –Intermediate: 4th Edition (TOLD-I:4; nine studies) is designed for children older than 8 years. The remaining eight assessments covered most of the 4-12 primary school age range selected for this study and included the following assessments: Assessment of Comprehension and Expression (ACE 6-11; seven studies), Comprehensive Assessment of Spoken Language (CASL; 12 studies), Clinical Evaluation of Language Fundamentals – 5th Edition (CELF-5; nine studies), Diagnostic Evaluation of Language Variance – Norm Referenced (DELV-NR; ten studies), Illinois Test of Psycholinguistic Abilities - 3rd Edition (ITPA-3; eight studies), Listening Comprehension Test – 2nd Edition (LCT-2; seven studies), Oral and Written Language Scales – 2nd Edition (OWLS-2; eight studies) and Woodcock Johnson 4th Edition Oral Language (WJIVOL; six studies). These 15 selected assessments are summarised in Table 2.5 with regards to author, publication date and language area assessed.

During the selection process, 61 assessments were excluded as not meeting the study criteria. These assessments are summarised in Table 2.6 with regards to author, publication date, language area assessed and reason for exclusion.

Table 2.5

*Summary of Assessments Included in the Review*

| Acronym and Name of Test (Authors; Publication date) | Age-group | Areas assessed<br>Subtests (norm-referenced)<br>Composite scores derived from subtests |
|---|---|---|
| **ACE6-11**<br>Assessment of Comprehension and Expression 6-11 (Adams, Cooke, Crutchley, Hesketh, & Reeves, 2001) | 6-11 years | **Spoken language including pragmatics.**<br>**Subtests:**<br>•Sentence comprehension, •Inferential comprehension, •Naming, •Syntactic formulation, •Semantic decisions, •Non-Literal comprehension, •Narrative propositions, •Narrative syntax/discourse<br>**Composite Scores:**<br>•Overall Language Score (Main Test or Extended version) |
| **ALL**<br>Assessment of Literacy and Language (Lombardino, Leiberman, & Brown, 2005) | [a] Pre-school - grade 1 | **Spoken and written language skills including phonemic awareness**<br>**Subtests:**<br>•Letter Knowledge, •Rhyme Knowledge, •Basic Concepts, •Receptive Vocabulary, •Parallel Sentence Production, •Ellison, •Word Relationships, •Rhyme Knowledge, •Phonics Knowledge, •Sound Categorisation, •Sight Word Recognition, •Listening Comprehension<br>**Composite Scores:**<br>•Emergent Literacy Index, •Language Index , •Phonological Index, •Phonological-Orthographic Index |
| **CASL**<br>Comprehensive Assessment of Spoken Language (Carrow-Woolfolk, 1999) | 3-21 years | **Spoken language including pragmatics**<br>**Subtests:**<br>•Comprehension of Basic Concepts, •Antonyms, •Synonyms, •Sentence Completion, •Idiomatic Language, •Syntax Construction, •Paragraph Comprehension of Syntax, •Grammatical Morphemes, •Sentence Comprehension of Syntax, •Grammaticality Judgement, •Non-Literal Language, •Meaning from Context, •Inference, •Ambiguous Sentences, •Pragmatic Judgment<br>**Composite Scores:**<br>•Core Language, •Lexical/Semantic (7;0 -21 years only), •Syntactic (7;0 -21 years only) •Supra-linguistic (7;0 -21 years only) , •Receptive Index (7;0-10;11 years only), •Expressive Index (7;0-10;11 years only) |
| **CELF-5**<br>Clinical Evaluations of Language Fundamentals – 5th Edition (Wiig, Semel, & Secord, 2013) | 5;0- 21;11 years | **Spoken language; supplemental tests for reading, writing and pragmatics**<br>**Subtests:**<br>•Sentence Comprehension, •Linguistic Concepts, •Word Structure, •Word Classes , •Following Directions , •Formulated Sentences, •Recalling Sentences, •Understanding Spoken Paragraphs, •Word Definitions, •Sentence Assembly, •Semantic Relationships, •Sentence Comprehension •Reading Comprehension (not used in composite scores), •Structured Writing (not used in composite scores) •Pragmatics profile (observational checklist, not used in composite scores)<br>**Composite Scores:**<br>•Core Language, •Receptive Language, •Expressive Language •Language Content, •Language Structure, •Language Memory |
| **CELF-P:2**<br>Clinical Evaluation of Language Fundamentals : Preschool – 2nd Edition (Wiig, Secord, & Semel, 2004) | 3;0 -6;11 years | **Spoken language**<br>**Subtests:**<br>•Sentence Structure, •Word Structure, •Expressive Vocabulary, •Concepts and Following Directions, •Recalling Sentences, •Basic Concepts, •Word Classes<br>**Composite Scores:**<br>•Core Language, •Receptive Language, •Expressive Language •Language Content, •Language Structure, •Recalling Sentences in Context (not used in composite scores), •Phonological Awareness (not used in composite scores), •Pre-Literacy Rating Scale (not used in composite scores) |
| **DELV-NR**<br>Diagnostic Evaluation of Language Variation – Norm referenced **(Seymour, W., & de Villiers, 2005)** | 4-9 years | **Spoken language:**<br>**Subtests:**<br>•Semantics, •Syntax, •Pragmatics, •Phonology (not used in composite score)<br>**Composite scores:**<br>•Total Language Score |

| **ITPA-3**<br>Illinois Test of Psycholinguistic Abilities – 3rd Edition<br>(Hammill, Mather, & R, 2001) | 5;0-12;11 years | **Spoken and written language:**<br>**Subtests:**<br>•Spoken Analogies, •Spoken Vocabulary, •Morphological Closure, •Syntactic Sentences, •Sound Deletion •Rhyming Sequences, •Sentence Sequencing, •Written Vocabulary, •Sight Decoding, •Sound Decoding, •Sight Spelling, •Sound Spelling<br>**Composite Scores:**<br>•General Language, •Spoken Language, •Written Language, •Semantics •Grammar, •Phonology, •Comprehension, •Word Identification, •Spelling, •Sight-Symbol Processing , •Sound-Symbol Processing |
|---|---|---|
| **LCT-2**<br>The Listening Comprehension Test—2nd Edition<br>(Bowers, Huisingh, & LoGiudice, 2006) | 6-11 years | **Spoken language**<br>**Subtests:**<br>•Main Idea, •Details •Reasoning, •Vocabulary, •Understanding Messages<br>**Composite Score:**<br>•Total Score |
| **NRDLS**<br>Reynell Developmental Language Scale -4th Edition<br>(Edwards, Letts, & Sinka, 2011) | 3;0 – 7;5 years | **Spoken language.**<br>**Subtests:**<br>•Comprehension, •Production<br>**Composite Scores:**<br>•Total Language Score |
| **OWLS-II**<br>Oral and Written Language Scales – 2nd Edition<br>(Carrow-Woolfolk, 2011) | 3-21 years | **Spoken language.**<br>**Subtests:**<br>•Listening Comprehension, •Oral Expression •Reading Comprehension, •Written Language<br>**Composite Scores:**<br>•Oral Language •Written Language, •Receptive Language •Expressive Language, •Overall Language |
| **PLS-5**<br>Preschool Language Scales-5th Edition<br>(Zimmerman, Steiner, & Pond, 2011) | Birth-7;11 years | **Spoken language**<br>**Subtests:**<br>•Auditory Comprehension, •Expressive Communication<br>**Composite scores:**<br>•Total Language Score |
| **TELD-3**<br>Test of Early Language Development – 3rd Edition<br>(Hresko, Reid, & Hammill, 1999) | 3,0- 7,11 | **Spoken language.**<br>**Subtests:**<br>•Receptive Language, •Expressive Language<br>**Composite Scores:**<br>•Spoken Language |
| **TOLD-I:4**<br>Test of Language Development – Intermediate: 4th Edition<br>(Newcomer & Hammill, 2008) | 8;0 – 17 years | **Spoken language.**<br>**Subtests:**<br>•Sentence Combining, •Picture Vocabulary, •Word Ordering, •Relational Vocabulary, •Morphological Comprehension, •Multiple Meanings •Word Discrimination (not used in composite scores), •Phonemic Analysis (not used in composite scores), •Word Articulation (not used in composite scores)<br>**Composite Scores:**<br>•Listening, •Organising, •Speaking •Grammar, •Semantics •Spoken Language |
| **TOLD-P:4**<br>Test of Language Development – Primary: 4th Edition<br>(Hammill & Newcomer, 2008) | 4;0 – 8;11 years | **Spoken language.**<br>**Subtests:**<br>•Sentence Combining, •Picture Vocabulary, •Word Ordering, •Relational Vocabulary, •Morphological Comprehension, •Multiple Meanings<br>**Composite Scores:**<br>•Listening, •Organising, •Speaking •Grammar, •Semantics •Spoken Language |
| **WJIVOL**<br>Woodcock Johnson IV Tests of Oral Language<br>(Shrank, Mather, & McGrew, 2014) | 2-90 years | **Spoken language**<br>**Subtests:**<br>•Picture Vocabulary, •Oral Comprehension •Segmentation, •Rapid Picture Naming, •Sentence Repetition, •Understanding Directions, •Sound Blending, •Retrieval Fluency, •Sound Awareness<br>**Composite Scores:** •Oral Language •Broad Oral Language •Oral Expression, •Listening Comprehension •Phonetic coding •Speed of Lexical Access |

Notes: [a] Normative data is based on United States school grade level. No normative data is provided for age level in this assessment.

Table 2.6

*Summary of Assessments Excluded from the Review*

| | Name of Test | Author and publication date | Age-group (years) | Language area/s tested | Reasons for exclusion |
|---|---|---|---|---|---|
| 1 | Adolescent Language Screening Test (ALST) | Morgan & Gillford (1984) | 11-17 | Pragmatics, receptive vocabulary, expressive vocabulary, sentence formulation, morphology and phonology | Not published within last 20 years |
| 2 | Aston Index Revised (Aston) | Newton & Thomson (1982) | 5-14 | Receptive language, written language, reading, visual perception, auditory discrimination | Not published within last 20 years |
| 3 | Bracken Basic Concept Test – Expressive (BBCS-E) | Bracken (2006) | 3-6;11 | Expressive: basic concepts | Not comprehensive language assessment |
| 4 | Bracken Basic Concept Test - 3rd Edition Receptive (BBCS:3-R) | Bracken (2006) | 3-6;11 | Receptive: basic concepts | Not comprehensive language assessment |
| 5 | Bankson Language Test-Second Edition (BLT-2) | Bankson (1990) | 3,0-6,11 | Semantics, syntax/morphology and pragmatics | Not published within last 20 years |
| 6 | Boehm Test of Basic concepts 3rd Edition (Boehm- 3) | Boehm (2000) | Grades K-2 (US) | Basic concepts | Not comprehensive language assessment |
| 7 | Boehm Test of Basic Concepts Preschool 3rd Edition (Boehm-3 Preschool) | Boehm (2001) | 3,0-5,11 | Relational concepts | Not comprehensive language assessment |
| 8 | British Vocabulary Scale - 3rd Edition (BPVS-3) | Dunn, Dunn & Styles (2009) | 3-16 | Receptive vocabulary | Not comprehensive language assessment |
| 9 | Clinical Evaluation of Language Fundamentals – 5th Edition Metalinguistic (CELF-5 Metalinguistic) | Wiig & Secord (2013) | 9;0 – 21;0 | Higher level language: making inferences, conversation skills, multiple meanings and figurative language | Not comprehensive language assessment |
| 10 | Clinical Evaluations of Language Fundamentals – 5th Edition Screening (CELF-5 Screening) | Semel, Wiig & Secord (2013) | 5;0- 21;11 | Receptive and expressive semantics and syntax | Screening assessment |
| 11 | Comprehensive Receptive and Expressive Vocabulary Test Second Edition (CREVT-3) | Wallace & Hammill (2013) | 5-89 | Receptive and expressive vocabulary | Not comprehensive language assessment |
| 12 | Compton Speech and Language Screening Evaluation – Revised Edition | Compton (1999) | 3-6 | Expressive and receptive language, articulation, auditory memory and oral-motor co-ordination | Screening Assessment |
| 13 | Executive Functions Test Elementary | Bowers & Huisingh (2014) | 7;0 – 12;11 | Higher level language: working memory, problem solving, inferring and making predictions | Not comprehensive language assessment |
| 14 | Expressive Language Test – 2nd Edition (ELT-2) | Bowers Huisingh, LoGiudice, Orman (2010) | 5;0-11;0 | Expressive language: sequencing, metalinguistics, grammar and syntax | Not comprehensive language assessment |
| 15 | Expressive One-Word Vocabulary Test - 4th Edition (EOWPVT-4) | Martin & Brownell (2011) | 2-80 | Expressive vocabulary (picture naming) | Not comprehensive language assessment |

| 16 | Expression, Reception and Recall of Narrative Instrument (ERRNI) | Bishop (2004) | 4 - 15 | Narrative skills: story comprehension and retell | Not comprehensive language assessment |
|---|---|---|---|---|---|
| 17 | Expressive Vocabulary Test – Second Edition (EVT-2) | Williams (2007) | 2;6–90+ | Expressive vocabulary and word retrieval | Not comprehensive language assessment |
| 18 | Fluharty Preschool Screening Test Second Edition (FPSLST-2) | Fluharty (2000) | 3;0-6;11 | Receptive and expressive language: sentence repetition, answering questions, describing actions, sequencing events and articulation. | Screening Assessment |
| 19 | Fullerton Language Test for Adolescent Second Edition (FLTA-2) | Thorum (1986) | 11- Adult | Receptive and expressive language | Not published within last 20 years |
| 20 | Grammar and Phonology Screening Test (GAPS) | Van der Lely (2007) | 3;5-6;5 | Grammar and pre reading skills | Not Comprehensive language assessment |
| 21 | Kaufman Survey of Early Academic and Language Skills (K-SEALS) | Kaufman & Kaufman (1993) | 3;0-6;11 | Expressive and receptive vocabulary, numerical skills and articulation | Not published in last 20 years |
| 22 | Kindergarten Language Screening Test, Second Edition (KLST-2) | Gauthier & Madison (1998) | 3;6-6;11 | General language: question comprehension, following commands, sentence repetition, comparing and contrasting objects and spontaneous speech | Screening Assessment |
| 23 | Language Processing Test 3 Elementary (LPT-3:P) | Richard & Hanner (2005) | 5-11 | Expressive semantics: word association, categorising words, identifying similarities between words, defining words, describing words | Not comprehensive language assessment |
| 24 | Montgomery Assessment of Vocabulary Acquisition (MAVA) | Montgomery (2008) | 3-12 | Receptive and expressive vocabulary | Not comprehensive language assessment |
| 25 | North Western Syntax Screening Test (NSST) | Lee (1969) | Unknown | Syntax and morphology | Not published in last 20 years |
| 26 | Peabody Picture Vocabulary test 4th Edition (PPVT-IV) | Dunn & Dunn (2007) | 2;6 – 90 | Receptive vocabulary | Not comprehensive language assessment |
| 27 | Pragmatic Language Skills (PLSI) | Gillam & Miller (2006) | 5;0-12;11 | Pragmatics | Not comprehensive language assessment |
| 28 | Preschool Language Assessment Instrument - Second Edition (PLAI-2) | Blank, Rose & Berlin (2003) | 3.0 - 5,11 | Discourse | Not comprehensive language assessment |
| 29 | Preschool Language Scales - 5th Edition Screener (PLS-5 Screener) | Zimmerman (2013) | Birth- 7;11 | General language | Screening assessment |
| 30 | Receptive One-Word Picture Vocabulary Tests, Fourth Edition (ROWPVT-4) | Martin & Brownell (2010) | 2;0- 70 | Receptive vocabulary | Not comprehensive language assessment |
| 31 | Renfrew Action Picture Test – Revised (RAPT- Revised) | Renfrew (2010) | 3-8 | Expressive language: information content, syntax and morphology | Not comprehensive language assessment |
| 32 | Renfrew Bus Story – Revised edition (RBS- Revised) | Renfrew (2010) | 3-8 | Narrative retell | Not comprehensive language assessment |
| 33 | Rhode Island Test of Language Structure | Engen & Engen (1983) | 3-6 | Receptive syntax (designed for hearing impairment but has norms for non-hearing impairment) | Not comprehensive language assessment |
| 34 | Screening Kit of Language Development (SKOLD) | Bliss &Allen (1983) | 2-5 | General language | Not published within last 20 years |

| 35 | Screening Test for Adolescent Language (STAL) | Prather & Breecher (1980) | 11-18 | General language | Not published in last 20 years |
| 36 | Social Emotional Evaluation (SEE) | Wiig (2008) | 6;0 -12;0 | Social skills and higher-level language | Not comprehensive language assessment |
| 37 | Social Language Development Test Elementary (SLDT-E) | Bowers, Huisingh, & LoGiudice (2008) | 6-11 | Language for social interaction | Not comprehensive language assessment |
| 38 | Structured Photographic Expressive Language Test - Third Edition (SPELT-3) | Dawson & Stout (2003) | 4,0 -9,11 | Expressive syntax and morphology | Not comprehensive language assessment |
| 39 | Structured Photographic Expressive Language Test Preschool 2nd Edition (SPELT-P:2) | Dawson et al (2005) | 3;0-5;11 | Expressive syntax and morphology | Not comprehensive language assessment |
| 40 | Test for Auditory Comprehension of Language–Fourth Edition (TACL-4) | Carrow-Woolfolk (2014) | 3,0-12,11 | Receptive vocabulary, syntax and morphology | Not comprehensive language assessment |
| 41 | Test of Auditory Reasoning and processing skills (TARPS) | Gardner (1993) | 5-13;11 | Auditory processing: verbal reasoning, inferences, problems solving, acquiring and organising information | Not published within last 20 years |
| 42 | Test for Examining Expressive Morphology (TEEM) | Shipley (1983) | 3;0 – 7;0 | Expressive morphology | Not published within last 20 years |
| 43 | Test of Grammatical Impairment (TEGI) | Rice & Wexler (2001) | 3;0–8;0 | Syntax and morphology | Not comprehensive language assessment |
| 44 | Test of Grammatical Impairment – Screener (TEGI -Screener) | Rice & Wexler (2001) | 3-6;11 | Syntax and morphology | Screening assessment |
| 45 | Test of Language Competence-Expanded (TLC-E) | Wiig & Secord (1989) | 5;0 – 18;0 | Semantics, syntax and pragmatics | Not published within last 20 years |
| 46 | Test of Narrative language (TNL) | Gillam & Pearson (2004) | 5;0 – 11;11 | Narrative retell | Not comprehensive language assessment |
| 47 | Test of Pragmatic Language (TOLP-2) | Terasaki & Gunn (2007) | 6;0-18;11 | Pragmatic skills | Not comprehensive language assessment |
| 48 | Test of Problem Solving 3 Elementary (TOPS-3-Elementary) | Bowers, Huisingh, & LoGiudice (2005) | | Language-based thinking | Not comprehensive language assessment |
| 49 | Test of Reception of Grammar (TROG-2) | Bishop (2003) | 4+ | Receptive grammar | Not comprehensive language assessment |
| 50 | Test of Semantic Skills - Intermediate (TOSS-I) | Huisingh, Bowers, LoGuidice, & Orman (2004) | 9-13 | Receptive and expressive semantics | Not comprehensive language assessment |
| 51 | Test of Semantic Skills – Primary (TOSS-P) | Bowers, Huisingh, LoGiudice, & Orman (2002) | 4-8 | Receptive and expressive semantics | Not comprehensive language assessment |
| 52 | Test of Word Finding—Second Edition (TWF-2) | German (2000) | 4;0 – 12;11 | Expressive vocabulary: word finding | Not comprehensive assessment |
| 53 | Test of Word Finding in Discourse (TWFD) | German (1991) | 6;6 -12;11 | Word finding in discourse | Not comprehensive assessment |
| 54 | Test of Word Knowledge (TOWK) | Wiig & Second (1992) | 5-17 | Receptive and expressive vocabulary | Not published within last 20 years |
| 55 | Token Test for Children – Second edition (TTFC2) | McGHee, Ehrler, & DiSimoni (2007) | 3;0 – 12;11 | Receptive: understanding of spoken directions | Not comprehensive language assessment |

| 56 | Wellcomm: A speech and language toolkit for the early years (Screening tool) English norms | Hurd, McQueen, & Sandwell Primary Care Trust (2010) | 6 months - 6 years | General language | Screening Assessment |
|----|------|------|------|------|------|
| 57 | Wh – question comprehension test | Vicker (2002) | 4 -Adult | Wh-question comprehension | Not comprehensive language assessment |
| 58 | Wiig Assessment of Basic Concepts (WABC) | Wiig (2004) | 2;6-7;11 | Receptive and expressive: basic concepts | Not comprehensive assessment |
| 59 | Word Finding Vocabulary Test – Revised Edition (WFVT) | Renfrew (2010) | 3-8 | Expressive vocabulary: word finding | Not comprehensive language assessment |
| 60 | The WORD Test 2 Elementary (WORD-2) | Bowers, Huisingh, LoGiudice, & Orman (2004) | 6-11 | Receptive and expressive vocabulary | Not comprehensive language assessment |
| 61 | Utah Test of Language Development (UTLD-4) | Mecham (2003) | 3,0 – 9,11 | Expressive semantics, syntax and morphology | Not comprehensive language assessment |

The seven identified articles were sourced from database searches and grey literature. These included studies investigating structural and convergent validity (hypothesis testing) of the CASL (Hoffman et al., 2011; Reichow, Salamack, Paul, Volkmar, & Klin, 2008), convergent validity (hypothesis testing) using the CELF-P:2 and the DELV-NR (Pesco & O'Neill, 2012) convergent validity (hypothesis testing) of the CELF-P:2 (Kaminski, Abbott, Aguayo, Latimer, & Good, 2014) convergent validity (hypothesis testing) of the TELD-3 (Spaulding, 2012), diagnostic accuracy of the CELF-P (Eadie et al., 2014), and internal consistency and test-retest reliability of the CASL pragmatic judgment subtest (McKown, Allen, Russo-Ponsaran, & Johnson, 2013). All articles appeared to have been published by authors independent of the developers of the assessments. The seven included articles are described in Table 2.7.

Table 2.7

*Articles Selected for Review*

| Author | Assessment | COSMIN property rated from study |
|--------|-----------|--------------------------------|
| Eadie et al. (2014) | CELF-P:2 (Australian) Diagnostic accuracy | Investigation of sensitivity and specificity of CELF:P-2 at age 4 years against Clinical Evaluation of Language Fundamentals – 4th Edition (CELF-4) at age 5 years |
| Hoffman et al. (2011) | CASL Structural Validity Hypothesis testing | Investigation of the construct (structural) validity of the CASL using factor analysis. Investigation of convergent validity between the CASL and Test of Language Development –Primary: 3rd Edition (TOLD-P:3) |
| Kaminski et al. (2014) | CELF-P:2 Hypothesis testing | Investigation of predictive validity and convergent validity between CELF:P-2 and Preschool Early Literacy Indicators (PELI) |
| McKown et al. (2013)* | CASL Internal consistency Reliability (test-retest) | Examination of the internal consistency of the Pragmatic Judgement subtest of the CASL* Examination of test-retest reliability of the Pragmatic Judgement subtest of the CASL |
| Pesco and O'Neill (2012) | CELF:P-2 DELV-NR Hypothesis testing | Investigation of performance on the DELV-NR and CELF-P:2 to be predicted by the Language Use Inventory (LUI) |
| Reichow et al. (2008) | CASL Hypothesis testing | Examination of the convergent validity between selected subtests from the CASL with the Vineland Adaptive Behaviour Scales. |
| Spaulding (2012) | TELD-3 Hypothesis testing | Investigation of consistency between severity classification on the TELD-3 and the Utah Test of Language Development – 4th Edition (UTLD-4) |

Notes: *This subtest forms part of the overall composite score on the CASL

The assessment manuals for all the selected assessments were not available through open sources and were only accessible by purchasing the assessment. Only three published articles by authors of assessments were identified. One of these contained information on the development, standardisation and psychometric properties of the NRDLS (Letts, Edward, Schaefer, & Sinka, 2014). This study was not included in this review as it was published after the assessment manual and contained no new information. Similarly, another article by the developers of the NRLDS (Letts, Edwards, Sinka, Schaefer, & Gibbons, 2013) examined the relationship between the NRDLS scores and economic status. This study was also reported in the manual and was not included. One other study by Seymour and colleagues (Seymour & Zurer-Pearson, 2004) described the rationale and proposed structure for the DELV-NR assessment; however, this study was also not included as it did not contain information on the psychometric properties of the final version of the assessment.

**2.4.2. Psychometric evaluation.** The results of the COSMIN ratings of the psychometric quality of the 15 assessments are listed in Table 2.8. Thirteen of the 15 assessment manuals included studies on the six COSMIN measurement properties evaluated in this review. One assessment (NRDLS) presented no examination of structural validity and another assessment (WJIVOL) did not have a reliability study using the subtests that primarily contribute to overall composite language scores. Manuals that contained studies with more than one reliability study i.e. inter-rater or test-retest reliability were given a rating for each type of reliability. Similarly, manuals with more than one study of hypothesis testing i.e. convergent or discriminant validity were given more than one rating for hypothesis testing. This is noted in Table 2.8 with two ratings for reliability and hypothesis testing where multiple studies were identified.

Ratings for each measurement property are shown as percentage of total points available and classified according to quartile in which percentage falls: Excellent (Excell) = 100-75.1, Good =75-50.1, Fair = 50-25.1 and Poor = 25-0. The rating of measurement properties based on percentages of all items allows for the overall quality of a study be considered, however it also means that it was possible for studies to be rated 'excellent' or 'good' overall when individual items may have been rated 'poor' for methodology. The footnotes in Table 2.8 indicate where studies were rated 'excellent', 'good' or 'fair' overall, but were identified as having a 'poor' rating for important items, such as: uni-dimensionality of the scale not checked prior to internal consistency calculation; sample size not stated or small; type of statistical analysis used unclear or inappropriate statistical analysis according to COSMIN; error measurement calculated using Cronbach's Alpha or split-half reliability method; time interval between assessment administrations not deemed appropriate; internal consistency calculated using split-half reliability; or correlations between subtests reported for structural validity rather than factor analysis.

Studies with COSMIN ratings of 'fair' or higher were then rated on the evidence provided in the study outcome for each measurement property using the criteria as summarised in Table 2.3. These results are reported in Table 2.8 underneath the methodological rating for each assessment. As COSMIN ratings represent the overall methodological quality of assessments and outcome ratings rate studies against specific methodological criteria, it is possible for studies with good COSMIN ratings to be rated as indeterminate for study outcome due to the presence of specific but significant flaws. The overall rating given after considering the methodological quality and outcome of all available studies (Table 2.8) is provided in Table 2.9.

Table 2.8

*Ratings of Methodological Quality and Study Outcome of Reliability and Validity Studies for Selected Assessments. Study Outcome Ratings are based on Terwee et al. (2007) and Schellingerhout et al. (2011).*

| Assessment | Manual or article | Internal Consistency | Reliability | Error Measurement | Content Validity | Structural Validity | Hypothesis Testing |
|---|---|---|---|---|---|---|---|
| ACE6-11 | ACE6-11 Manual | 77.8 [a] Excell ? | Test-retest 75.9 Excell ? | 53.3 [d] Good ? | 42.9 Fair ? | 25[h] Poor NE | Convergent 52.2 Good + Discriminant 23.5 Poor NE |
| ALL | ALL Manual | 75.0 [b] Good ? | Test-retest 72.4 Good ? inter-rater 50 [c] Fair ? | 20 [d] Poor NE | 92.9 Excell + | 33.3 [b] Fair ? | Convergent 52.2 Good + Discriminant 52.9 Good + |
| CASL | CASL Manual | 57.1 [g] Good ? | Test-retest 56.0 [e] Good ? | 40 [d] Fair ? | 71.4 Good ? | 33.3 [b] Fair ? | Convergent 39.1 Fair + Discriminant 58.8 Good + |
| | Hoffman et al., 2011 | NR | NR | NR | NR | 33.3 [b] Fair ? | Convergent 73.9 Good ± |
| | McKown et al., 2013 | 83.3 [a] Excell ? | Test-retest 62.0 [e] Good ? | NR | NR | NR | NR |
| | Reichow et al., 2008 | NR | NR | NR | NR | NR | Convergent 52.2 Good ? |

| Assessment | Manual or article | Internal Consistency | Reliability | Error Measurement | Content Validity | Structural Validity | Hypothesis Testing |
|---|---|---|---|---|---|---|---|
| CELF-5 | CELF-5 Manual | 71.4 [g] Good ? | Test-retest 72.4 Good ? <br> Inter-rater 66.7 Good + | 40 [d] Fair ? | 71.4 Good + | 58.3 Good ? | Convergent 65.2 Good + <br> Discriminant 52.9 Good + |
| CELF:P-2 | CELF:P-2 Manual | 71.4 [b] Good ? | Test-retest 72.4 Good ? <br> Inter-rater 50.0 [c] Fair ? | 40 [d] Fair ? | 64.3 Good + | 33.3 [b] Fair ? | Convergent 47.8 Fair + <br> Discriminant 58.8 Good + |
| | Kaminski et al., 2014 | NR | NR | NR | NR | NR | Convergent 56.5 Good ± |
| | Pesco & O'Neill, 2012 | NR | NR | NR | NR | NR | Convergent 47.8 Good ± |
| | * Manual for ALL | NR | NR | NR | NR | NR | Convergent 65.2 Good + |
| | * Manual for PLS-5 | NR | NR | NR | NR | NR | Convergent 69.6 Good + |
| DELV-NR | DELV-NR Manual | 66.7 [a] Good ? | Test-retest 69 Good ? <br> Inter-rater 50 [c] Fair ? | 40 [d] Fair ? | 57.1 Good ? | 50 [h] Fair ? | Convergent 34.8 Fair ? <br> Discriminant 41.2 Fair ? |
| | * Pesco & O'Neill, 2012 | NR | NR | NR | NR | NR | Convergent 47.8 Good ± |
| ITPA-3 | ITPA-3 Manual | 71.4 [b] Good ? | Test-retest 62.1 Good ? <br> Inter-rater 41.7 Fair ? | 40 [d] Fair ? | 57.1 Fair ? | 50 Fair ? | Convergent 34.7 Fair + <br> Discriminant 41.2 Fair ? |
| LCT-2 | LCT-2 Manual | 50 [a] Fair ? | Test-retest 34.6 Fair ? <br> Inter-rater 25 [c] Poor NE | 40 [d] Fair ? | 28.5 Fair ? | 50 [h] Fair ? | Discriminant 29.4 [f] Fair + |
| NRDLS | NRDLS Manual | 66.7 [a] Good ? | Test-retest 60.0 Good ? | 40.0 [d] Fair ? | 57.1 Good ? | NR | Convergent 52.2 Good + <br> Discriminant 35.3 Fair + |
| OWLS-II | OWLS-II Manual | 57.1 [g] Good ? | Test-retest 72.4 Good ? <br> Inter-rater 50 Fair + | 40 [d] Fair ? | 71.4 Good ? | 33.4 [b] Fair ? | Convergent 21.7 Poor NR <br> Discriminant 47.1 Fair + |
| PLS-5 | PLS-5 Manual | 50 [g] Fair ? | Test-retest 69.0 Good ? <br> Inter-rater 50 [g] Fair ? | 40 [d] Fair ? | 71.4 Good ? | 57.1 [h] Good ? | Convergent 56.5 Good + <br> Discriminant 52.9 Good + |
| TELD-3 | TELD-3 Manual | 61.1 [a] Good ? | Test-retest 72.4 Good ? <br> Inter-rater 33.3 [g] Fair ? | 33.4 [d] Fair ? | 71.4 Good ? | 41.7 [h] Fair ? | Convergent 39.1 Fair ? <br> Discriminant 35.3 Fair + |
| | Spaulding, 2012 | NR | NR | NR | NR | NR | Convergent 47.8 Fair ? |

| Assessment | Manual or article | Internal Consistency | Reliability | Error Measurement | Content Validity | Structural Validity | Hypothesis Testing |
|---|---|---|---|---|---|---|---|
| TOLD-I:4 | TOLD-P:4 Manual | 71.4 [b] Good ? | Test-retest 72.4 Good ? Inter-rater 41.7 [c] Fair ? | 40 [d] Fair ? | 57.1 Fair ? | 33.4 [b] Fair ? | Convergent 60.9 Good + Discriminant 35.3 Fair ? |
| TOLD-P:4 | TOLD-I:4 Manual | 71.4 [b] Good ? | Test-retest 69.0 Good ? Inter-rater 50 Fair ? | 40 [d] Fair ? | 57.1 Fair ? | 50 [b] Fair ? | Convergent 60.9 Good + Discriminant 35.3 Fair + |
| WJIVOL | WJIVOL Manual | 57.2 [g] Good ? | NE | 40 [d] Fair ? | 78.6 Excell ? | 50 [b] Fair ? | Convergent 43.5 Fair + Discriminant 41.2 Fair ? |

Notes: Excellent (Excell) = 100-75.1, Good =75-50.1, Fair = 50-25.1 and Poor = 25-0; NR = No study reported for this measurement property in this publication; NE = study not evaluated as 'poor' methodological rating;

+, ?, - = See Table 2.3; [a] Uni-dimensionality of scale not checked prior to internal consistency calculation; [b] Sample size for factor analysis not stated or small; [c] Type of statistical analysis used unclear or inappropriate statistical analysis according to COSMIN; [d] Error measurement calculated using Cronbach alpha or split-half reliability

Table 2.9

*Level of Evidence for Each Assessment*

| Assessment | Internal Consistency | Reliability | Error Measurement | Content Validity | Structural Validity | Hypothesis Testing |
|---|---|---|---|---|---|---|
| ACE6-11 | ? | ? | ? | ? | ? | ++ |
| ALL | ? | ? | ? | +++ | ? | +++ |
| CASL | ? | ? | ? | ? | ? | ++* |
| CELF-5 | ? | ++ | ? | ++ | ? | +++ |
| CELF:P-2 | ? | ? | ? | ++ | ? | +++ * |
| DELV-NR | ? | ? | ? | ? | ? | ? * |
| ITPA-3 | ? | ? | ? | ? | ? | + |
| LCT-2 | ? | ? | ? | ? | ? | + |
| NRDLS | ? | ? | ? | ? | NA | ++ |
| OWLS-II | ? | + | ? | ? | ? | + |
| PLS-5 | ? | ? | ? | ++ | ? | +++ |
| TELD-3 | ? | ? | ? | ? | ? | + |
| TOLD-I:4 | ? | ? | ? | ? | ? | ++ |
| TOLD-P:4 | ? | ? | ? | ? | ? | ++ |
| WJIVOL | ? | NA | ? | ? | ? | + |

Notes:  +++ or --- Strong evidence positive/negative result; ++ or -- Moderate evidence positive/negative result; + or - Limited evidence positive/negative result; ± Conflicting evidence across different studies; ?, Unknown due to poor methodological quality; NA, no information available; Blue shading, positive evidence; yellow shading, evidence unknown.

*Some studies outside of the manuals were rated as having conflicting evidence within the same study.

For seven assessments, studies examining diagnostic accuracy were identified. This information came from the respective manuals and one article. Data on sensitivity, specificity, positive predictive power and negative predictive power for these seven assessments are presented in Table 2.10. With regards to the assessments reviewed in this study, sensitivity and specificity indicates the percentage of children with language impairment identified by the assessment as having language impairment (sensitivity) and the percentage of children with no language impairment identified as having no language impairment (specificity). Higher values indicate higher diagnostic accuracy, with literature suggesting that values between 90-100% (0.90-1.00) indicate 'good' accuracy and values between 80-89% (0.80-0.89) indicate 'fair' accuracy (Greenslade, Plante, & Vance, 2009; Plante & Vance, 1994). Predictive power indicates how precise an assessment is in predicting children with language impairment (Positive Predictive Power or PPP) and children without language impairment (Negative Predictive Power or NPP) for different cut-off scores against a pre-determined prevalence base rate. Higher predictive values indicate better precision in predictive power. It should be noted that whilst these results from diagnostic accuracy studies are reported without being rated for methodological quality, significant methodological concerns were noted and are reported in the discussion section of this study.

Table 2.10

*Diagnostic Accuracy Data Reported for Each Assessment*

| Assessment | Manual or article | Criterions | Sensitivity % | Specificity % | PPP % | NPP% |
|---|---|---|---|---|---|---|
| **ALL** | ALL Manual | 10% base rate for population sample; 50, 70, 80 & 90% base rate for referral population; Other criterion not specified | -1SD =98<br>-1.5SD = 86<br>-2SD = 54 | -1SD = 89<br>-1.5SD = 96<br>-2SD = 98 | 10% base rate:<br>-1SD = 50<br>-1.5SD = 73<br>-2SD = 77<br>80% base rate:<br>-1SD = 97<br>-1.5SD = 99<br>-2SD = 99 | 10% base rate:<br>-1SD = 100<br>-1.5SD = 98<br>-2SD = 95<br>80% base rate:<br>-1SD = 93<br>-1.5SD = 30<br>-2SD = 35 |
| **CELF-5** | CELF-5 Manual | 10% base rate for population sample; 50, 60, 70 & 80% base rate for referral population; Other criterion not specified | -1SD=100<br>-1.3SD=97<br>-1.5SD=85<br>-2SD=57 | -SD=91<br>-1.3SD=97<br>-1.5SD=99<br>-2SD=100 | 10% base rate:<br>-1SD = 55<br>-1.3SD = 78<br>-1.5SD = 86<br>-2SD = 100<br>80% base rate:<br>-1SD = 98<br>-1.3SD = 99<br>-1.5SD = 100<br>-2SD = 100 | 10% base rate:<br>-1SD = 100<br>-1.3SD = 100<br>-1.5SD = 98<br>-2SD = 95<br>80% base rate:<br>-1SD = 100<br>-1.3SD = 89<br>-1.5SD = 62<br>-2SD = 37 |
| **CELF:P-2** | CELF:P-2 Manual | 20% base rate for population sample; 50, 70, 80, 90% for referral sample | NR | NR | 20% base rate:<br>-1SD = 53<br>-1.5SD = 66<br>-2SD =82<br>80% base rate:<br>-1SD = 95<br>-1.5SD = 97<br>-2SD =99 | 20%base rate:<br>-1SD = 95<br>-1.5SD = 91<br>-2SD = 86<br>80%base rate:<br>-1SD = 57<br>-1.5SD = 39<br>-2SD = 28 |
| | Eadie et al., 2013 | CELF-P:2 scores at 4 years against CELF-4 scores at 5 years | -1.25SD=64.0<br>-2SD=42.1 | -1.25SD =92.9<br>-2SD = 98.6 | NR | NR |
| **DELV-NR** | DELV-NR Manual | 10% base rate for population sample; 50, 60, 70 & 80% base rate for referral population; Other criterion not specified | -1SD =95<br>-1.5SD = 69<br>-2SD = 36 | -1SD = 93<br>-1.5SD = 99<br>-2SD = 100 | 10% base rate:<br>-1SD = 61<br>-1.5SD = 87<br>-2SD = 100<br>80% base rate:<br>-1SD = 98<br>-1.5SD = 100<br>-2SD = 100 | 10% base rate:<br>-1SD = 99<br>-1.5SD = 97<br>-2SD = 93<br>80% base rate:<br>-1SD = 84<br>-1.5SD = 45<br>-2SD = 28 |
| **PLS-5** | PLS-5 Manual | 20% base rate for population sample; 50, 70, 80, 90% for referral sample; Other criterion not specified | With standard score 85 as cut-off = 91 | With standard score 85 as cut-off = 78 | 20% base rate:<br>-1SD = 51<br>-1.5SD = 73<br>-2SD = 78<br>80% base rate:<br>-1SD = 94<br>-1.5SD = 98<br>-2SD = 98 | 20% base rate:<br>-1SD = 95<br>-1.5SD = 92<br>-2SD = 87<br>80% base rate:<br>-1SD = 55<br>-1.5SD = 41<br>-2SD = 30 |

| Assessment | Manual or article | Criterions | Sensitivity % | Specificity % | PPP % | NPP% |
|---|---|---|---|---|---|---|
| **TOLD-I:4** | TOLD-I:4 Manual | Criterion against other assessments: [a] PLOS, [b] PPVT-3, [c] TOLD-P:4, [d] WISC-IV & [e] Global Language score; Other criterion not specified | With Standard Score 90 as cut-off: [e] Global Language Score = 77 | With Standard Score 90 as cut-off: [e] Global Language Score = 89 | With Standard Score 90 as cut-off: [e] Global Language Score = 71 | NR |
| **TOLD-P:4** | TOLD-P:4 Manual | Criterion against other assessments: [a] PLOS, [f] TOLD-P:4 & [g] Global Language Score; Other criterion not specified | With Standard Score 90 as cut-off: [g] Global Language Score = 75 | With Standard Score 90 as cut-off: [g] Global Language Score = 87 | With Standard Score 90 as cut-off: [g] Global Language Score = 71 | NR |

Notes: PPP = Positive Predictive Power; NPP = Negative Predictive Power; Base rate for population sample = percentage of population expected to identify with language impairment; Base rate for referral population = percentage of children referred for assessment who identify with language impairment; NR= Not reported in this study; SD= Number of standard deviations selected as cut-off for calculation; [a] PLOS= Pragmatic Language Observation Scale; [b] PPVT-3 = Peabody Picture Vocabulary Test - Third Edition; [c] TOLD-P:4 = Test of Oral Language Development – Primary: 4th Edition; [d] WISC-IV = Weschler Intelligence Scale for Children – 4th Edition (Verbal Comprehension Composite); [e] Global Language Score = Metavariable combining PLOS, PPVT-3, TOLD-P:4, WISC-IV scores`; [f] TOLD-P:4= Test of Language Development-Intermediate: 4th Edition; [g] Global Language Score= Metavariable combining PLOS & TOLD-P:4 scores.

## 2.5. Discussion

**2.5.1. Methodological quality of studies**. In this study, a total of 121 studies across all six measurement properties were rated for methodological quality. Of these, 5 were rated as 'excellent' for overall methodological quality, 55 rated as 'good', 56 rated as 'fair' and 5 rated as 'poor'. However, whilst almost half (n= 60) of all studies rated as 'good' or better overall, only one quarter (n=29) of all studies had sufficient methodological quality to meet the criteria in Table 2.3 based on a revision of criteria proposed by Terwee et al. (2007) and Schellingerhout et al. (2011). Therefore, over half of the studies with generally good design were identified as having specific weaknesses which ultimately compromised the usefulness of findings. Methodological flaws in studies examining psychometric quality of language

assessments have also been noted in other literature (LEADERS, 2014, 2015). Therefore, there is a great need for improvements in the design and reporting of studies examining psychometric quality of language assessments for children. Clinicians and researchers also need to be critical of methodology when viewing the results of studies examining reliability and validity of assessments.

Overall, across all measurement properties, reporting on missing data was insufficient, with few studies providing information on the percentage of missing items or a clear description of how missing data was handled. Bias may be introduced if missing data is not determined as being random (Bennett, 2011); therefore, this information is important when reporting on the methodology of studies examining psychometric quality.

A lack of clarity in reporting of statistical analysis was also noted, with a number of assessment manuals not clearly reporting the statistics used. For example, studies used terms such as 'correlation' or 'coefficient' without specifying the statistical procedure used in calculations. Where factor analysis or intra-class correlations were applied in structural validity or reliability studies, few studies reported details such as the rotational method or formula used. Lack of clear reporting creates difficulty for independent reviewers and clinicians to appraise and compare the quality of evidence presented in studies.

COSMIN ratings for *internal consistency* were rated between 'excellent' and 'fair' with most rated as 'good'. However, only two thirds of the reviewed assessments used the statistical analysis required for evidence of internal consistency according to Terwee et al. (2007) and Schellingerhout et al. (2011); that is, Cronbach's Alpha or Kuder-Richardson Formula–20. The remaining assessments (CASL, CELF-5, OWLS-II, PLS-5 and WJIVOL) used a split-half reliability method. Of the ten studies that utilised Cronbach alpha, five studies did not have uni-dimensionality of the scale confirmed through factor analysis and the remaining five did not have an adequate sample size. For internal consistency results to have

interpretable meaning, the scale needs to be identified as being uni-dimensional (Terwee et al., 2012).

With regards to *reliability* most assessments rated in the range of 'good' or 'fair'. Three assessments (ACE6-11, CASL and NRDLS) reported test-retest reliability but did not examine inter-rater reliability. One assessment (WJIVOL) did not present with any reliability studies for the subtests that contribute to composite scores that target oral language. All other assessments included examinations of both test-retest and inter-rater reliability within the manuals. Two assessments (OWLS-II and TELD-3) were designed with alternate record forms and, although not included in this review, it was noted that these assessments also reported on the parallel-forms reliability. However, only two assessments (CELF-5 and OWLS-II) used the statistical analysis identified as optimal in Table 2.3, intra-class correlation or weighted kappa; and were thus the only two studies identified as having evidence of reliability.

COSMIN ratings for *measurement error* were rated the lowest of all measurement properties, with no studies rating better than 'fair'. All studies were rated 'poor' for statistical analysis as reliabilities calculated from split-half or Cronbach alpha were used to calculate standard error of measurement, which does not meet COSMIN's requirement of two administrations for evaluating measurement error (Terwee et al., 2012). Measurement error is the variability of random error that may affect assessment results. This is used to develop confidence intervals for scores and reflects the precision to which assessment scores for individuals can be reported.

Ratings for *content validity* varied considerably across different assessments. While most assessments mapped content onto modalities of comprehension and production and domains of semantics, syntax/morphology, pragmatics and phonology, different theoretical constructs were used to guide content selection. As no empirical evidence currently exists

regarding the modalities or domains of language that should be assessed and the criteria for determining impairment (Eadie et al., 2014; Tomblin et al., 1996; Tomblin & Zhang, 2006; Van Weerdenburg, Verhoeven, & Van Balkom, 2006), assessments that rated lower were those that did not: 1) provide a clear definition of theoretical construct, 2) provide a clear rationale for how items were selected for the purpose of the assessment or 3) have an assessment of content from experts during the development of the assessment. The assessments identified as having evidence of content validity were the ALL, CELF-5, CELF:P-2 and PLS-5.

COSMIN ratings for *structural validity* studies rated between 'good' and 'poor'. Of the 15 assessments rated, nine assessments (ALL, CELF-5, CELF-P:2, ITPA-3, CASL, OWLS-II, TOLD-P:4, TOLD-I:4, WJIVOL) had an examination of structural validity using factor analysis which is the statistical method required for evidence of structural validity according to COSMIN and Schellingerhout et al. (2011). However, of these nine assessments, only two (CELF-5 and ITPA-3) were rated as 'good' or 'excellent' for the sample size used. Sample size for factor analysis depends on the number of items in an assessment. As comprehensive language assessments tend to have a large number of items, many studies did not have sample sizes large enough for an 'excellent' factor analysis rating on COSMIN, despite the sample appearing large. No studies reported on the percentage of explained variance in structural validity studies, therefore no studies were rated as having a good level of evidence in this measurement property.

Five assessment manuals (ACE6-11, DELV-NR, LCT-2, PLS-5 and TELD-3) did not report on a structural validity study using factor analysis but reported a study measuring correlations between subtests; however, this is not sufficient evidence of structural validity according to COSMIN. One assessment (NRDLS) did not provide any evidence to support structural validity through either factor analysis or an examination of correlations between

subtests. Structural validity studies are important to examine the extent to which an assessment reflects the underlying constructs being measured in both the overall score and the subtests.

The majority of studies relating to *hypothesis testing* rated as 'fair' or 'good' for overall methodological quality. All 15 assessments reported on a comparison between the performance of children with language impairment and typical children and all, except the LCT-2, provided information on convergent validity with related measures of language. Fourteen studies presented with some level of evidence in this measurement property, with only one study (DELV-NR) lacking in studies with sufficient methodological quality for evidence to be determined. For three assessments (CASL, CELF-P, DELV-NR) convergent validity studies outside of the manuals presented with conflicting results. However, it should be noted that these assessments were three of the very few assessments for which independent studies were identified. As such, the possibility exists that conflicting evidence may appear for other assessments if independent studies were available.

Studies on *diagnostic accuracy* were available for half of the selected assessments. This information included studies examining positive predictive power (PPP) using estimates of the percentage of children expected to have language impairment in a sample population and studies examining sensitivity and specificity using another assessment as a criterion. Population estimates were set at 10-20% for an overall child population and 60-90% for a population of children referred to services for assessment. Many studies also included PPP calculations with a base percentage of 50%. Most assessments presented data using a range of different standard deviations as cut-off points (between 1 standard deviation and 2 standard deviations) for identification of impairment. The variation in population estimates and cut-off points may reflect the lack of consistency with criteria for diagnosis of language impairment

which is noted in literature (Greenslade et al., 2009; Spaulding et al., 2006; Tomblin et al., 1996).

Diagnostic accuracy studies were not rated for methodological quality; however significant methodological flaws were noted in the reporting of information. The evaluated article (Eadie et al., 2014) reported the sample size and sample selection methods used in the study, however no manuals reported this information. When this information is lacking, it is impossible for speech pathologists to evaluate the quality of study or to determine if the sample population represents the clinical population for which the assessment is to be used (Dollaghan & Horner, 2011). Of the studies reporting on sensitivity and specificity against another criteria for identifying language impairments, only the TOLD-P:4 manual, TOLD-I:4 manual and the article (Eadie et al., 2014) provided any description of the reference measure used and time length between assessment administrations. This lack of reporting is a serious flaw as it does not allow for the impact of potential classification errors by the reference standard to be considered in evaluating the validity of findings (Betz et al., 2013; Dollaghan & Horner, 2011). When the reference standard is not specified it also creates difficulty when attempting to compare findings for different assessments or compare different studies for the same assessment. Therefore, evidence regarding the diagnostic accuracy of currently available language assessments is lacking due to an overall trend with poor methodological quality. Improvements in methodological quality and reporting of studies are needed to provide this evidence and to assist Speech Pathologists in understanding the diagnostic utility of available assessments (Dollaghan & Horner, 2011; LEADERS, 2014, 2015).

An important discovery was that all the studies examined in this review used statistical methods solely from classical test theory (CTT), as opposed to item response theory (IRT). Although some manuals made reference to the use of IRT methods in the initial development of assessment items, no studies reported any details or outcomes for these

methods. Whilst COSMIN does not currently indicate a preference between these two methods, IRT methods are increasingly being utilised for the development of assessments within fields such as psychology and have numerous reported advantages over CTT-only methods (Edelen & Reeve, 2007; Reise, Ainsworth, & Haviland, 2005). Further investigation is needed to examine reasons for the lack of IRT methods in the development of child language assessments.

      **2.5.2. Comparison between manuals and independent studies**. Comparisons between manuals and independent articles are limited to instances where studies with adequate methodology from both a manual and an article are available for a measurement property. These included three instances examining convergent validity of the CASL, CELF:P-2 and DELV-NR (Hoffman et al., 2011; Kaminski et al., 2014; Pesco & O'Neill, 2012). In all three of these examples, the articles were rated as reporting conflicting evidence whilst ratings in manuals were rated as having positive evidence. Pesco and O'Neill (2012) examined the ability for DELV-NR and CELF:P-2 scores to be predicted by earlier scores on another assessment, the Language use Inventory (LUI). The study reported correlations above the 0.5 suggested by Schellingerhout et al. (2011) for one of five age groups investigated, although the authors named a significant correlation for three age groups. Kaminski et al. (2014) examined predictive and convergent validity between the CELF-P:2 and an assessment called the Preschool Early Literacy Indicators (PELI). In this study, correlations between composite scores were found to be slightly above the level suggested by Schellingerhout et al. (2011) for predictive validity and slightly below for convergent validity. Another study by Hoffman et al. (2011) examined convergent validity between the CASL and the Test of Language Development – Primary: 3rd Edition (TOLD-I:3). This study identified a correlation using Pearson's r above the level described as acceptable by Schellingerhout et al. (2011); however, further analysis using a t-test for significance

identified a significant difference between composite scores of the assessments. From this, the authors suggested that it may not be accurate to assume that the two different assessments can be used inter-changeably with the same results.

The correlations reported in the CELF-P:2 manual (Wiig et al., 2004) for convergent validity were higher than the correlations reported in articles, however in the manual, the CELF-P:2 was compared to different versions of itself (CELF-P and CELF-4) and with a similar test published by the same publisher (PLS-4). Therefore, the correlations would be expected to be higher than the correlations reported in the article where the CELF-P:2 was compared to a language assessment with a different theoretical background. The time period between administrations of assessments also differed between studies, which may be a source of difference, given the potential for possible change in status of children over time.

The study by Hoffman et al. (2011) also examined structural validity of the CASL using factor analysis. Although this study was not identified as having adequate methodology due to small sample size, the results are interesting to note because different findings were reported in comparison to the factor analysis reported in the CASL manual (Carrow-Woolfolk, 1999). Hoffman et al. (2011) reported evidence of a single factor model however the manual reported a 3-factor model. However, the 3-factor model was only reported in the manual for children seven years and older, with a single factor model reported for ages six and below. The sample in the article included 6, 7 and 8 year-olds, therefore encompassing both these age-ranges. Furthermore, the two studies did not administer the same subtests from the CASL and both studies received a 'poor' COSMIN rating for sample size. Factor analysis on five subtests of the CASL collectively containing 260 items would require a sample size of over 1,300 for a COSMIN rating higher than 'poor'. Both these studies had sample sizes less than 250. Given the shortcomings of these studies, further studies with good methodology are required to provide evidence of structural validity.

Collectively, these findings indicate that further independent studies are required to examine the validity of different comprehensive language assessments for children. Further research is also required to determine if children are categorised similarly across different assessments with regards to diagnosis and severity of language impairment (Hoffman et al., 2011; Spaulding et al., 2012).

**2.5.3. Overall quality of language assessments**. It is acknowledged that speech pathologists should consider a range of factors as well as psychometric quality when selecting an assessment for use including the clinical population for which the assessment will be used, the purpose for which the assessment will be used and theoretical construct of the assessment (Bishop & McDonald, 2009). This study examined the reliability and validity of currently available assessments and identified that all assessments present with notable shortcomings when rated against methodological quality (COSMIN) and the criteria of evaluating findings of studies (Table 2.3). However, considering the data that is available, some assessments have more psychometric evidence to support use as diagnostic assessments. These assessments include: ALL, CELF-5, CELF:P-2 and PLS-5. It is noted that the ALL currently only provides grade level normative data for the United States of America population. The ALL, CELF-5 and PLS-5 were all rated as having 'strong' or 'moderate' evidence across two or more measurement properties. The CELF:P-2 was identified as having evidence in two measurement properties from the manual, however there was some conflicting information regarding hypothesis testing in independent literature. The ALL, CELF-5 and PLS-5 were not examined in independent literature. The DELV-NR, ITPA-3, LCT-2, TELD-3 and WJIVOL had no more than limited evidence for one measurement property. However, it should be noted that where evidence is reported as lacking, it does not mean that these assessments are not valid or reliable, but rather that further research is required to determine psychometric quality.

**2.5.4. Implications.** Standardized assessments are frequently used to make important diagnostic and management decisions for children with language impairment in both clinical and research contexts. For accurate diagnosis and provision of effective intervention, it is important that assessments chosen for use have evidence of good psychometric quality (Friberg, 2010). However, a previous study identified that speech pathologists may not be selecting child language assessments based on the psychometric quality reported in assessment manuals (Betz et al., 2013). Therefore, emphasis needs to be placed on the selection of assessments that are evidence-based and appropriate to the needs of the client, the speech pathologist and the service delivery context. Speech pathologists also need to advocate for improvements to the quality of both currently used assessments and those developed in the future.

This review also identifies areas in need of further research with regards to individual assessments and development of the field of child language assessment in general. Where an assessment does not present with an 'excellent' or 'good' level of evidence for all measurement properties, further research is required to determine if this evidence exists. In general, further information is particularly needed to provide evidence of structural validity, measurement error and diagnostic accuracy. The use of IRT methods for statistical analysis of psychometric properties of also identified as an area in need of further exploration within the field of child language assessment.

Very limited evidence of psychometric quality currently exists outside of what is reported in manuals for child language assessments and where evidence does exist, it does not always support information reported in manuals Assessment manuals are produced by developers who have commercial interest in the assessment. Furthermore, the reporting of psychometric quality in manuals is not peer-reviewed and can only be viewed after purchasing. When assessment developers make information on psychometric properties

available online or in published peer-reviewed journals, transparency is achieved and clinicians and researchers are able to review psychometric properties prior to purchasing assessments. A need for independent studies is also identified in order to provide additional information to data provided in assessment manuals. When information is able to be collated from a variety of different studies, then the evidence regarding psychometric quality of assessments will become more substantial.

This review identified a number of assessments that currently present with better evidence of psychometric quality than others, although substantially more data is required to show that any assessments have 'good' evidence. Until further information becomes available, it is suggested that speech pathologists favour assessments with better evidence when assessing the language abilities of school-aged children, provided that the normative sample is appropriate for the population in which the assessment is to be used. However, given that all assessments have limitations, speech pathologists should avoid relying on the results of a single assessment. Standardized assessment results should be supplemented with information from other assessment approaches (e.g., response to intervention, curriculum-based assessment, language sampling, dynamic assessment) when making judgements regarding diagnosis and intervention needs (Eadie et al., 2014; Hoffman et al., 2011). In addition, as it is possible that differences in underlying constructs between assessments contributes to differences in diagnostic abilities of assessments (Hoffman et al., 2011), it is important for speech pathologists to consider theoretical construct when choosing standardized assessments for use or when comparing results between different assessments.

**2.5.5. Limitations.** Due to a need to restrict size, responsiveness was not investigated in this review. It was, however, noted that no assessment manuals reported on responsiveness studies. These studies have a longitudinal design with multiple administrations of the assessment across time to measure sensitivity to change in a person's status. Evidence of

responsiveness is particularly important when assessments are to be used for measuring intervention outcomes or monitoring stability over time (Eadie et al., 2014; Polit, 2015). Therefore, further research is recommended to investigate the evidence for using comprehensive language assessments for these purposes. Further investigation is also needed to compare assessments across different English-speaking countries and cultural groups.

This review was confined to school-age language assessments that cover both the production and comprehension of spoken language. While this reflects current literature and clinical practice (Tomblin et al., 1996; Wiig, 2011), there may be clinical applications for assessments specific to one modality, for example when assessing language abilities of children who are non-verbal or have unintelligible speech. Assessments targeting single aspects of language such as semantics or syntax were also not included in this study, however, these may be used by Speech Pathologists (Betz et al., 2013), therefore an examination of psychometric quality of these assessments is recommended.

There is a need for future research to examine the psychometric quality of assessments for children who are bi-lingual or speaking English as a second language (Gillam et al., 2013). An examination of standardized written language assessments is also needed as there is a strong overlap between spoken and written language impairment in school-aged children (Bishop & Snowling, 2004; Snowling & Hulme, 2012). In addition, there is also a need for investigation into assessments that target activity and participation levels of the World Health Organisation's International Classification of Functioning and Disability – Child and Youth (McLeod & Threats, 2008; Roulstone, 2012).

## 2.6. Conclusion

This systematic review examines the psychometric quality of 15 currently available standardized language assessments for children aged 4-12 years. Overall, limitations were noted with the methodology of studies reporting on psychometric quality, indicating a great

need for improvements in the design and reporting of studies examining psychometric quality of both existing assessments and those that are developed in the future. As information on psychometric properties is primarily provided by assessment developers in manuals, further research is also recommended to provide independent evidence for psychometric quality. Whilst all assessments were identified as having notable limitations, four assessments: ALL, CELF-5, CELF:P-2 and PLS-5 were identified as currently having better evidence of reliability and validity. These four assessments are suggested for diagnostic use, provided they suit the purpose of the assessment process and are appropriate for the population being assessed. Emphasis on the psychometric quality of assessments is important for speech pathologists to make evidence-based decisions about the assessments they select when assessing the language abilities of school-aged children.

**References for Chapter 2**

Adams, C., Cooke, R., Crutchley, A., Hesketh, A., & Reeves, D. (2001). *Assessment of Comprehension and Expression 6-11*. London, U.K.: GL Assessment.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)*. Washington, DC: American Psychiatric Association.

Andersson, L. (2005). Determining the adequacy of tests of children's language. *Communication Disorders Quarterly, 26*(4), 207-225. doi:10.1177/15257401050260040301

Bennett, D. A. (2011). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health, 25*(5), 464-469. doi:10.1111/j.1467-842X.2001.tb00294.x

Betz, S. K., Eickhoff, J. K., & Sullivan, S. F. (2013). Factors influencing the selection of standardized tests for the diagnosis of Specific Language Impairment. *Language, Speech and Hearing Services in Schools, 44*, 133-143. doi:10.1044/0161-1461(2012/12-0093)

Bishop, D. V. M. (2011). Ten questions about terminology for 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., & McDonald, D. (2009). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of Language & Communication Disorders, 44*(5), 600-615. doi:10.1080/13682820802259662

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific langauge impairment: same or different? *Psychological Bulletin, 130*, 858-886. doi:10.1037/0033-2909.130.6.858

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS ONE, 11*(7), e0158753. doi:10.1371/journal.pone.0158753

Bowers, L., Huisingh, R., & LoGiudice, C. (2006). *The Listening Comprehension Test* (2 ed.). Austin, TX: Pro-Ed.

Caesar, L. G., & Kohler, P. D. (2009). Tools clinicians use: A survey of language assessment procedures used by school-based speech-pathologists. *Communication Disorders Quarterly, 30*(4), 226-236. doi:10.1177/1525740108326334

Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language*. Torrence, CA: WPS.

Carrow-Woolfolk, E. (2011). *Oral and Written Language Scales - Second Edition* (2 ed.). Bloomington, MN: Pearson.

Catts, H. W., Bridges, M. S., Little, T. D., & Tomblin, J. B. (2008). Reading achievement growth in children with language impairments. *Journal of Speech, Language, and Hearing Research, 51*, 1569-1579. doi:10.1044/1092-4388(2008/07-0259)

Conti-Ramsden, G., & Botting, N. (2004). Social difficulties and victimization in children with SLI at 11 years of age. *Journal of Speech, Language, and Hearing Research, 47*(1), 145-161. doi:doi:10.1044/1092-4388(2004/013)

Conti-Ramsden, G., Durkin, K., Simkin, Z., & Knox, E. (2009). Specific language impairment and school outcomes: Identifying and explaining variability at the end of compulsory education. *International Journal of Speech-Language Pathology, 44*(1), 15-35. doi:10.1080/13682820801921601

Cordier, R., Speyer, R., Chen, Y., Wilkes-Gillan, S., Brown, T., Bourke-Taylor, H., . . . Leicht, A. (2015). Evaluating the Psychometric Quality of Social Skills Measures: A Systematic Review. *PLoS ONE, 10*(7), e0132299. doi:10.1371/journal.pone.0132299

Dockrell, J. E., & Lindsay, G. (1998). The ways in which speech and language difficulties impact on children's access to the curriculum. *Child Language Teaching and Therapy, 14*(2), 117-133. doi:10.1177/026565909801400201

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. *Child and Adolescent Mental Health, 20*(2), 116-125. doi:10.1111/camh.12072

Dollaghan, C. A., & Horner, E. A. (2011). Bilingual language assessment: A meta-analysis of diagnostic accuracy. *Journal of Speech, Language and Hearing Research (Online), 54*(4), 1077-1088. doi:10.1044/1092-4388(2010/10-0093)

Eadie, P., Nguyen, C., Carlin, J., Bavin, E., Bretherton, L., & Reilly, S. (2014). Stability of language performance at 4 and 5 years: measurement and participant variability. *International Journal of Language and Communication Disorders, 49*(2), 215-227. doi:10.1111/1460-6984.12065

Ebbels, S. (2014). Introducing the SLI debate. *International Journal of Language & Communication Disorders, 49*(4), 377-380. doi:10.1111/1460-6984.12119

Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research, 16*(5), 5–18. doi:10.1007/s11136-007-9198-0

Edwards, S., Letts, C., & Sinka, I. (2011). *New Reynell Developmental Language Scales*. London, England: GL Assessment.

Fraser, J., & Conti-Ramsden, G. (2008). Contribution of phonological and broader language skills to literacy. *International Journal of Language and Communication Disorders, 43*(5), 552-569. doi:10.1080/13682820701778069

Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact of diagnostic decisions? *Child Language Teaching and Therapy, 26*(1), 77-92. doi:10.1177/0265659009349972

Gillam, R. B., Peña, E. D., Bedore, L. M., Bohman, T. M., & Mendez-Perezb, A. (2013). Identification of Specific Language Impairment in Bilingual Children: Assessment in English. *Journal of Speech, Language, and Hearing Research, 56*, 1813-1823. doi:10.1044/1092-4388(2013/12-0056)

Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the Structured Photographic Expressive Language Test—Preschool. *Language, Speech, and Hearing Services in Schools, 40*(2), 150-160. doi:10.1044/0161-1461(2008/07-0049)

Hammill, D. D., Mather, R., & R, R. (2001). *Illinois Test of Psycholinguistic Abilities* (3 ed.). Austin,TX, U.S.: Pro-Ed.

Hammill, D. D., & Newcomer, P. L. (2008). *Test of Oral Language Development - Primary: 4th Edition*. Austin, TX: Pro-Ed.

Harrison, L. J., McLeod, S., Berthelsen, D., & Walker, S. (2009). Literacy, numeracy, and learning in school-aged children identified as having speech and language impairment in early childhood. *International Journal of Speech-Language Pathology, 11*(5), 392-403. doi:10.1080/17549500903093749

Haynes, W. O., & Pindzola, R. H. (2012). *Diagnosis and Evaluation in Speech Pathology* (8th ed.). Harlow: Pearson Education Inc.

Hegde, M. N., & Pomaville, F. (2013). *Assessment of Communication Disorders in Children* (2 ed.). San Diego, CA: Plural publishing.

Hoffman, l. M., Leob, D. F., Brandel, J., & Gillam, R. B. (2011). Concurrent and construct validity of oral language measures with school-aged chidldren with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 20*, 1597-1608. doi:10.1044/1092-4388(2011/10-0213)

Hresko, W. P., Reid, D. K., & Hammill, D. D. (1999). *Test of Early Language Development* (3 ed.). Austin, TX: Pro-Ed.

Kaderavek, J. N. (2011). *Language Disorders in Children: Fundamental Concepts of Assessment and Intervention*. Upper Saddle River, NJ: Pearson Education Inc.

Kaminski, R. A., Abbott, M., Aguayo, K. B., Latimer, R., & Good, R. H. (2014). The preschool early literacy indicators: Validity and benchmark goals. *Topics in Early Childhood Special Education, 34*(2), 71-82. doi:10.1177/0271121414527003

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language and Communication Disorders, 35*(2), 165-188. doi:10.1080/136828200247133

LEADERS. (2014). Test Review: Clinical Evaluation of Language Fundamentals – Fifth Edition (CELF-5). Retrieved from https://edblogs.columbia.edu/leadersproject/files/2014/02/CELF5-Test-Review-LEADERS-1-29ay7q2.pdf

LEADERS. (2015). Test Review: Preschool Langage Scales -5th Edition (PLS-5). Retrieved from https://edblogs.columbia.edu/leadersproject/files/2015/05/PLS5-English-finaldraft-2dgptop.pdf

Leonard, L. B. (2009). Is Expressive Language Disorder an Accurate Diagnostic Category? *American Journal of Speech - Language Pathology, 18*(2), 115-123. doi:doi:10.1044/1058-0360(2008/08-0064)

Letts, C., Edward, S., Schaefer, B., & Sinka, I. (2014). The New Reynell Developmental Language Scales: Descriptive account and illustrative case study. *Child Language Teaching and Therapy, 30*(1), 103-116. doi:10.1177/0265659013492784

Letts, C., Edwards, S., Sinka, I., Schaefer, B., & Gibbons, W. (2013). Socio-economic status and language acquisition: children's performance on the new Reynell Developmental Language Scales. *International Journal Language Communication Disorders, 48*(2), 131-143. doi:0.1111/1460-6984.12004

Lindsay, G., Dockrell, J., Desforges, M., Law, J., & Peacey, N. (2010). Meeting the needs of children and young people with speech, language and communication difficulties. *International Journal of Speech-Language Pathology, 45*(4), 448-460. doi:10.3109/13682820903165693

Lombardino, L. J., Leiberman, R., & Brown, J. C. (2005). *Assessment of Literacy and Language*. San Antonio, TX: Pearson Psychcorp.

McCauley, R. J., & Swisher, L. (1984). Psychometric reveiw of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders, 49*, 34-42. doi:10.1044/jshd.4901.34

McCormack, J., Harrison, L. J., McLeod, S., & McAllister, L. (2011). A nationally representative study of the association between communication impairment at 4–5 years and children's life activities at 7–9 years *Journal of Speech, Language, and Hearing Research, 54*, 1328-1348. doi:10.1044/1092-4388(2011/10-0155)

McKown, C., Allen, A. M., Russo-Ponsaran, N. M., & Johnson, J. K. (2013). Direct Assessment of Children's Social-Emotional Comprehension. *Psychological Assessment, 25*(4), 1154–1166. doi:10.1037/a0033435

McLeod, S., & Threats, T. T. (2008). The ICF-CY and children with communication disabilities. *International Journal of Speech-Language Pathology, 10*, 92–109. doi:10.1080/1754950070183469

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med 6(6): e1000097*. doi:0.1371/journal.pmed1000097

Mokkink, L. B., Prinsen, C. A. C., Bouter, L. M., De Vet, H. C. W., & Terwee, C. B. (2015). The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Brazilian Journal of Physical Therapy, 20*(2), 105-113. doi:10.1590/bjpt-rbf.2014.0143

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., . . . de Vet, H. C. W. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology, 10*, 22-30. doi:10.1186/1471-2288-10-22

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology and definitions of measurement properties for health-related patient-reported outcomes. *Quality of Life Research, 19*, 539-549. doi:10.1016/j.jclinepi.2010.02.006

Mokkink, L. B., Terwee, C. B., Stratford, P. W., Alonso, J., Patrick, D. L., Riphagen, I., . . . de Vet, H. C. W. (2009). Evaluation of the methodological quality of systematic

reviews of health status measurement instruments. *Quality of Life Research, 18*(3), 313-333. doi:10.1007/s11136-009-9451-9

Newcomer, P. L., & Hammill, D. D. (2008). *Test of Language Development - Intermediate: 4th Edition* (4 ed.). Austin, TX, U.S.: Pro-Ed.

Paul, R., & Norbury, C. F. (2012a). Chapter 1: Models of child language disorders. In R. Paul & C. F. Norbury (Eds.), *Language Disorders from Infancy through Adolescence: Listening, Speaking, Reading, Writting and Communicating* (4 ed., pp. 1-21). St Louis, MI: Mosby Elsevier.

Paul, R., & Norbury, C. F. (2012b). Chapter 2: Assessment. In R. Paul & C. F. Norbury (Eds.), *Language disorders from infancy through adolescence: Listening, speaking, reading, writing and communicating* (4th ed., pp. 22-60). St Louis, MI: Mosby Elsevier.

Paul, R., & Norbury, C. F. (2012c). Chapter 8: Assessment of developing Language. In R. Paul & C. F. Norbury (Eds.), *Language Disorders from Infancy through Adolescence: Listening, Speaking, Reading, Writting and Communicating* (4 ed., pp. 280-346). St Louis, MI: Mosby Elsevier.

Pesco, D., & O'Neill, D. K. (2012). Predicting later language outcomes from the Language Use Inventory. *Journal of Speech, Language, and Hearing Research, 55*, 421–434. doi:10.1044/1092-4388(2011/10-0273)

Plante, E., & Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language, Speech, and Hearing Services in Schools, 25*(1), 15-24. doi:10.1044/0161-1461.2501.15

Polit, D. F. (2015). Assessing measurement in health: Beyond reliability and validity. *International Journal of Nursing Studies, 52*, 1746-1753. doi:10.1016/j.ijnurstu.2015.07.002

Reed, V. A. (2005). Assessment. In V. A. Reed (Ed.), *An Introduction to Children with Language Disorders* (2 ed., pp. 404–442). Boston, MA: Pearson Education.

Reichow, B., Salamack, S., Paul, R., Volkmar, F. R., & Klin, A. (2008). Pragmatic Assessment in Autism Spectrum Disorders: A Comparison of a Standard Measure With Parent Report. *Communication Disorders Quarterly, 29*(3), 169-176. doi:10.1177/1525740108318697

Reise, S. P., Ainsworth, A. T., & Haviland, M. G. (2005). Item Response Theory: Fundamentals, Applications, and Promise in Psychological Research. *Current directions in psychological science, 14*(2), 95-101. doi:10.1111/j.0963-7214.2005.00342.x

Roulstone, S., Coad, J., Ayre, A., Hambly, H., & Lindsay, G. (2012). The preferred outcomes of children with speech, language and communication needs and their parents. Retrieved from https://dera.ioe.ac.uk/16329/7/DFE-RR247-BCRP12_Redacted.pdf

Schellingerhout, J. M., Heymans, M. W., Verhagen, A. P., de Vet, H. C. K., B. W., & Terwee, C. B. (2011). Measurement properties of translated versions of neck-specific questionaires: a systematic review. *BMC Medical Research Methodology, 11*(87), 659–670. doi:10.1186/1471-2288-11-87

Seymour, H. N., W., R. T., & de Villiers, J. (2005). *Diagnostic Evaluation of Langauge Variation*. Bloomington, MN: Pearson Psychcorp.

Seymour, H. N., & Zurer-Pearson, B. (2004). Steps in designing and implementing an innovative assessment instrument. *Seminars in Speech and Language, 25*(1), 27-31. doi:10.1055/s-2004-824823

Shipley, G. S., & McAfee, J. G. (2009). *Assessment in Speech-Language Pathology: A Resource Manual* (4th ed.). Clifton Park, NY: Cengage Learning.

Shrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Oral language*. Rolling Meadows, IL: Riverside.

Snowling, M. J., & Hulme, C. (2012). Annual Research Reveiw: The nature and classification of redaing disorders - acommentary on proposals for DSM-5. *Journal of Child Psychology and Psychiatry, 53*, 593-607.

Spaulding, T. J. (2012). Comparison of severity ratings on norm-referenced tests for children with specific language impairment. *Journal of Communication Disorders, 45*, 59-68.

Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Elligibility criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech and Hearing Services in Schools, 37*, 61-72. doi:10.1044/0161-1461(2006/007)

Spaulding, T. J., Szulga, M. S., & Figueroa, C. (2012). Using norm-referenced tests to determine severity of language impairment in children: Disconnect between US policy makers and test developers. *Language, Speech, and Hearing Services in Schools, 43*(2), 176-190. doi:10.1044/0161-1461(2011/10-0103)

Speyer, R., Cordier, R., Kertscher, B., & Heijnen, B. (2014). Psychometric Properties of Questionnaires on Functional Health Status in Oropharyngeal Dysphagia: A Systematic Literature Review. *BioMed Research International*. doi:10.1155/2014/458678

Terwee, C. B., Bot, S. D. M., deBoer, M. R., van der Windt, D. A. W. M., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionaires. *Journal of Clinical Epidemiology, 60*, 34-42. doi:10.1016/j.jclinepi.2006.03.012

Terwee, C. B., de Vet, H. C. W., & Mokkink, L. B. (2015). Reply to 'COSMIN for quality rating systematic reviews on psychometric properties'. *Physical Therapy Reviews, 20*(4), 266-267. doi:DOI 10.1179/1743288X15Y.0000000021

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & de Vet, H. C. W. (2012). Rating the methological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research, 21*, 651-657. doi:10.1007/s11136-011-9960-1

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 39*(6), 1284-1294. doi:10.1044/jshr.3906.1284

Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research, 49*(6), 1193-1208. doi:10.1044/1092-4388(2006/086)

Uijen, A. A., Heinst, C. W., Schellevis, F. G., van den Bosch, W. J., van de Laar, F. A., Terwee, C. B., & Schers, H. J. (2012). Measurement properties of questionnaires measuring continuity of care: a systematic review. *PLoS ONE, 7*(7), e42256. doi:10.1371/journal.pone.0042256

Van Weerdenburg, M., Verhoeven, L., & Van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry, 47*(2), 176–189. doi:10.1111/j.1469-7610.2005.01454.x

Vrijman, C., Homan, M. L., Limpens, J., van der Veen, W., Wolkerstorfer, A., Terwee, C. B., & Spuls, P. I. (2012). Measurement Properties of Outcome Measures for Vitiligo: A Systematic Review. *Archives of Dermatology, 148*(11), 1302-1309. doi:10.1001/archdermatol.2012.3065

Wiig, E. H. (2011). How SLD manifests in oral expression and reading comprehension. In D. P. A. Flanagan, V.C (Ed.), *Essentials of Specific Learning Disbility Identification*. Hoboken, NJ: John Wiley and Sons.

Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals: Preschool - 2nd Edition* (2 ed.). Bloomington, MN: Pearson.

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals - 5th Edition*. Bloomington, MN: Pearson.

World Health Organisation. (2015). *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (Vol. 10). Geneva, Switzerland: World Health Organisation.

Yew, S. G. K., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry, 54*(5), 516-524. doi:10.1111/jcpp.12009

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales-5th Edition* (5 ed.). Bloomington, MN: Pearson.

**Chapter 3.**

**Describing Language Assessments for School-Aged Children: A Delphi Study Overview for Chapter 3 (Journal Article 2)**

Chapter 3 relates to research area two. This chapter describes the findings from a Delphi study conducted to obtain consensus on a taxonomy of terminology for describing language assessments. Terminology from the taxonomy was then utilised in the large national survey presented in thesis Chapters 5 and 6. In addition to supporting survey research, the taxonomy provides professional terminology that may be used in SLP professional training and future research studies.

This chapter contains an Accepted Manuscript of an article published by Taylor & Francis in the International Journal of Speech-Language Pathology on 11 January 2019, available online: https://doi.org/10.1080/17549507.2018.1552716. The spelling and wording contained within this chapter is that of the published manuscript.

Note that the taxonomy presented in this chapter was developed through a Delphi study (lead by expert opinion) and not a literature review. Delphi study participants were provided with an initial document outlining the key features by which assessments are distinguished within the extant literature and were provided with a proposed structure for organising these distinctions (see Supplementary Appendix 3.1. for references to how the taxonomy links to current knowledge and literature in the field). The purpose of this initial document was to orient participants to the study aim and provide a common basis for discussion. The Delphi study process then allowed for participants to suggest distinctions or terms that should be added, removed or changed based on literature, clinical practice or expert opinion (see Supplementary Appendix 3.2 for an outline of how the taxonomy developed across the three rounds).

Copies of the Delphi study survey questions are contained in Supplementary

Appendices 3.3, 3.4.and 3.5. Copies of the case studies used in Delphi rounds two and three

are contained in Supplementary Appendices 3.6 and 3.7.

# Describing language assessments for school-aged children: A Delphi study

Deborah Denman[1*], Jae-Hyun Kim[2, 1], Natalie Munro [3,1], Renée Speyer [4, 1, 5], Reinie Cordier[1]

[1] Faculty of Health Sciences, Curtin University, Perth, Australia

[2] Department of Linguistics, Macquarie University, Sydney, Australia

[3] Faculty of Health Sciences, The University of Sydney, Sydney, Australia

[4] Department Special Needs Education, University of Oslo, Oslo, Norway

[5] Department of Otorhinolaryngology and Head and Neck Surgery, Leiden University Medical Centre, Leiden, The Netherlands

*Corresponding author: E-mail: deborah.denman@postgrad.curtin.edu.au

Short title: Describing Language Assessments for Children

## 3.1. Abstract

Purpose: Given the barriers that inconsistent terminology poses for the Speech-Language Pathology (SLP) profession, this study aimed to develop an agreed upon taxonomy with well-defined categories for describing language assessment practices for children.

Method: A taxonomy with illustrative terms for describing assessments across four aspects (modality/domain, purpose, delivery, and form) was developed with reference to contemporary literature. In a three round Delphi study, SLPs with expertise in child language were asked to indicate their level of agreement with the taxonomy and provide feedback. Participants were also asked to apply the taxonomy by categorising assessments presented in case studies.

Results: A total of 55 participants completed round one, while 43 and 32 completed rounds two and three respectively. Agreed consensus with the taxonomy was achieved in both rounds one and two, with at least 88% of participants agreeing with each aspect and 100% agreeing with the overall structure. In round three, agreement was reached on 7/10 components for one case study and 4/10 for the other.

Conclusion: The development of this taxonomy represents a significant step towards providing detailed terminology for describing language assessments. Future research is needed to investigate implementation strategies to facilitate consistent application of the taxonomy by SLPs.

**3.2. Background**

Internationally, the problem of inconsistent use of professional terminology by speech language pathologists (SLPs) is widely recognised (Walsh & IGOTF-CSD., 2006). Often one term may be used to refer to a range of different concepts or, conversely, different concepts are described by the same term (Walsh, 2005). Inconsistently applied terminology leads to breakdowns in professional communication and thus limits scientific debate needed to advance the profession. Lack of detailed terminology also hinders research translation as practices may not be described well enough to be replicated (Roulstone, 2015; Walsh & IGOTF-CSD., 2006).

In the field of child language assessment, many models and terms exist for describing the different types of language assessments that SLPs may use. As a result, the way in which one SLP conceptualises and describes their language assessment may well be different to another SLP's description of the same assessment. This creates significant barriers for the collection of accurate data on current practice both within and across service agencies (Cowie et al., 2001). Without an accurate understanding of current SLP assessment practice, it is difficult to compare current practice with evidence-based practice and thus identify clinical recommendations that align contemporary practice with policy and research evidence (Eadie, 2003).

A framework frequently used to describe SLP assessment practice is the International Classification of Functioning and Disability (ICF) (McLeod & Threats, 2008; World Health Organisation, 2015). This framework was designed to provide a structure by which concepts related to health and well-being may be viewed but, as such, lacks detail for describing language assessment (Barnes & Bloch, 2018; Hughes & Orange, 2007). Since communication spans all aspects of health and well-being, it is acknowledged that SLPs often experience difficulty mapping assessment practices onto ICF categories (Barnes & Bloch, 2018; Hughes & Orange, 2007). Considerable disparity exists across literature with regards to

how language assessments are classified within the ICF. For example, in some studies, norm-referenced language assessments, such as editions of the Clinical Evaluation of Language Fundamentals (Wiig, Semel, & Secord, 2004) are identified as assessing the body structure and functioning category of the ICF, while other studies identify these measures as assessing the activity category of the ICF (Paul & Norbury, 2012; Westby & Washington, 2017). The development of frameworks that are specifically targeted at describing SLP practices may facilitate greater consistency with regards to how assessments are described and thus enhance professional communication (Barnes & Bloch, 2018).

Specifically, within the field of paediatric SLP, there are a number of terms used to describe language assessments. One common feature is to describe the skill domain targeted in the assessment. This may be through the use of Bloom and Lahey's taxonomy, which describes language domains across three aspects including form, content and use (Lahey, 1988); or through terms such as semantics, syntax, morphology, narrative or executive functioning (Larson & McKinley, 2003; Paul & Norbury, 2012). A second way in which assessments may be described is according to the purpose of the assessment. Categories include analytical or prognostic; summative or formative; or distinctions related to diagnostic purposes, screening, selecting intervention or determining service eligibility (Dockrell & Marshall, 2015; Newton, 2007; Paul & Norbury, 2012; Wade, 2004). Assessments may also be described by the method in which the assessment is conducted or the environmental context targeted in the assessment. Examples of methods include parent questionnaires, tests administered either face-to-face or via telehealth, or assessments conducted by automated computer software. Examples of terms related to environmental context include curriculum-based, naturalistic or authentic (Parsons, Law, & Gascoigne, 2005; Paul & Norbury, 2012). Finally, assessments may be identified by the type of data collected or the type of tasks embedded in the assessment. This includes terms such as norm-referenced, criterion referenced and dynamic for describing type of data collected; or terms such as discrete-skill,

functional, contextualised or language sampling for describing the types of tasks being assessed (American Speech and Hearing Association, 2018; Kaderavek, 2015).

In addition to the presence of numerous sets of terms for conceptualising language assessments, the definitions of these terms are often not precisely defined or are defined differently across literature. For example, the term standardised has been used to describe any assessment that has structured guidelines for administration (regardless of the type of data collected), but has also been used interchangeably with the term norm-referenced to refer specifically to assessments that provide normative data from a sample of age-matched peers (Kaderavek, 2015). Terminology used to describe assessments that are non-standardised in nature is even more loosely defined, with terms such as authentic, alternative, informal, naturalistic, behavioural and observational all used with unrestrained boundaries for the types of assessments covered by these terms (Caesar & Kohler, 2009; Hegde & Pomaville, 2017). Furthermore, detail in assessment practice is not captured through the use of one framework or one set of terms. Two assessments described by the same term could be vastly different. For instance, a morphology assessment could refer to a series of clinician directed sentence completion tasks organised developmentally or an analysis of the morphological forms produced in a language sample (Paul & Norbury, 2012). Similarly, a language sample might be a highly structured, norm-referenced narrative retell task or observations by an SLP during unstructured free play (Westerveld & Claessen, 2014). To describe assessments in detail, SLPs need access to a framework that facilitates the conceptualisation of language assessments across multiple distinguishing features.

Given the current problems associated with terminology, there is a pressing need for actions that facilitate rigour and consistency with regards to the terms SLPs use for describing child language assessment (Walsh & IGOTF-CSD., 2006). It is evident that a single framework or a list of terms is unlikely to solve all problems related to such a complex problem (Walsh, 2005). Nonetheless, solutions are needed for situations where terminology

must be consistently applied in order to be useful, such as when collecting survey data on the types of practices SLPs use (Cowie et al., 2001). With this in mind, the establishment of an agreed-upon taxonomy for conceptualising various child language assessment practices is a logical step towards addressing some of the challenges associated with inconsistent terminology. In addition to facilitating data collection, such a taxonomy has the potential to stimulate much-needed professional discussion and reflection on assessment practice, which is vital for continued advancement in the field (Eadie, 2003; Roulstone, 2001). There is also a significant need for further research examining the application of professional terminology. This will assist in better understanding the issues and complexities associated with developing consistent use of terminology in the SLP field (Walsh & IGOTF-CSD., 2006).

**3.2.1 The current study.** This study employed a Delphi study technique to address two aims: (1) to develop a taxonomy (i.e. categorisation framework) that is agreed upon by experts and provides distinct, well-defined categories for describing language assessment practises for children aged 4-18 years; and (2) to examine SLP application of a taxonomy for describing language assessments in clinical contexts. For the purposes of this study, *language assessment* may be any data-gathering action including case histories, test performance, language sampling, behavioural observations, reports from significant others, and reports on educational achievement (Paul & Norbury, 2012). Given this is the first study to examine terminology for describing SLP assessment practice and acknowledging the potential complexities associated with developing consensus, the participant group in this study focussed on SLPs in a single country (Australia). It is envisaged that outcomes from this initial study will then act as the basis for further future research internationally.

**3.3. Method**

This study used a Delphi study technique with mixed-methods data collection and analysis (Tapio, Paloniemi, Varho, & Vinnari, 2011). The Delphi technique is a structured process which aims to develop group consensus on a defined topic through a series of survey

rounds (Boulkedid, Abdoul, Loustau, Sibony, & Alberti, 2011). The same participants complete each round (though not all may continue with each round) and rounds are held until consensus is reached (or it becomes apparent that consensus cannot be reached). Results from previous rounds are used to inform changes that may facilitate consensus in subsequent rounds.

**3.3.1. Participants.** Criteria for participation in the Delphi study were: (1) eligibility for certified practicing membership with the Australian professional association for SLPs (Speech Pathology Australia); and (2) having spent at least 5 years (full-time equivalent) in the last 10 years engaged in professional activities, where 50% or more of professional time is related to children aged 4-18 years with a language support needs. These activities may include: research, academic teaching, consultancy, resource development, provision of SLP professional development, provision of clinical services or a combination of these activities. Potential participants were contacted by email after being identified from the Speech Pathology Australia Find a Speech Pathologist website, the 2016 Speech Pathology Australia National Conference attendance contact list, and from the professional networks of the authors. In some states, recruited participants were also asked to identify other potential participants.

A total of 202 invitations were emailed and all SLPs who responded to invitations were sent a link to the first survey. As each survey was developed based upon the results of the preceding Delphi round, participants who did not complete a survey round were excluded from subsequent rounds. This helped to ensure that all participants had the same knowledge of the taxonomy. The number of participants who completed each Delphi round was 55 in round one (71.4% response rate), 43 in round two (78.2% response rate) and 32 in round three (74.4% response rate). Participant demographics for each round are presented in Table 3.1.

Table 3.1

*Demographics of Participants in the Delphi Study*

| Category | Subcategory | Round One n (%) | Round Two n (%) | Round Three n (%) |
|---|---|---|---|---|
| **State** | QLD | 7 (12.7%) | 7 (16.3%) | 5 (15.6%) |
| | NSW | 10 (18.2%) | 7 (16.3%) | 6 (18.8%) |
| | ACT | 1 (1.8%) | 1 (2.3%) | 0 (0.0%) |
| | VIC | 16 (29.1%) | 11 (25.6%) | 9 (28.1%) |
| | TAS | 3 (5.5%) | 2 (4.7%) | 2 (6.3%) |
| | NT | 3 (5.5%) | 1 (2.3%) | 1 (3.1%) |
| | SA | 7 (12.7%) | 7 (16.3%) | 6 (18.8%) |
| | WA | 8 (14.5%) | 7 (16.3%) | 3 (9.4%) |
| | Total | 55 | 43 | 32 |
| **Current Employment*** | Health Sector (government or non-government) | 5 (9.1%) | 5 (11.6%) | 2 (6.3%) |
| | Education Sector (government or non-government) | 18 (32.7%) | 17 (39.5%) | 16 (50.0%) |
| | Private Practice/Small Business | 10 (18.2%) | 7 (16.3%) | 4 (12.5%) |
| | University | 13 (23.6%) | 10 (23.3%) | 7 (21.9%) |
| | Other agency (government or non-government) | 3 (5.5%) | 2 (4.7%) | 0 (0.0%) |
| | Currently not working as SLP | 1 (1.8%) | 0 (0.0% | 0 (0.0%) |
| | Work across two of the above sectors | 5 (9.1%) | 2 (4.7%) | 3 (9.4%) |
| | Total | 55 | 43 | 32 |
| **Qualifications in addition to Bachelor or Graduate Equivalent degree*** | Masters or PhD | 24 (43.6%) | 18 (41.9%) | 15 (46.8%) |
| | Diploma (Education or Psychology) | 2 (3.6%) | 2 (4.7%) | 1 (3.1%) |
| | No other qualifications | 29 (52.7%) | 23 (53.5%) | 16 (50.0%) |
| | Total | 55 | 43 | 32 |
| **Years of experience (Full-time equivalent)*** | 5-10 years | 10 (18.2%) | 7 (16.3%) | 5 (15.6%) |
| | 11-15 years | 10 (18.2%) | 9 (20.9%) | 8 (25.5%) |
| | 16-21 years | 13 (23.6%) | 9 (20.9%) | 9 (28.1%) |
| | 21-30 years | 12 (21.8%) | 9 (20.9%) | 5 (15.6%) |
| | 30+ years | 10 (18.2%) | 9 (20.9%) | 5 (15.6%) |
| | Total | 55 | 43 | 32 |

*Note.* *As reported by participant

**3.3.2. Procedure.** Following a deductive (top-down) approach (DeJong, Horn, Gassaway, Slavin, & Dijkers, 2004), theoretical literature and research publications related to language assessment for school-aged children were reviewed by the first author to identify key concepts and terms that provide both a description of and differentiation between assessments. These concepts and terms were further developed through discussions within the research team and organised into an initial taxonomy. This taxonomy consisted of four aspects for describing features of assessments including: language modalities and domains, purposes, delivery methods and contexts, and the assessment form in terms of type of data collected and type of tasks used. The initial taxonomy was then presented to study participants in a three round Delphi study for feedback. Each Delphi round was conducted as an online survey using Qualtrics software. The round one survey was piloted initially with two SLPs to check clarity of questions and completion time before being opened to Delphi study participants. Each survey was estimated as taking 90 minutes to complete. Delphi rounds were conducted between April-October 2017 with each survey being accessible for three to seven weeks. The study details were outlined at the beginning of each survey; participants were required to indicate consent to participate before accessing the remainder of the survey content. The study was ethically approved by the Curtin University Human Research Ethics Committee (Approval number: HRE2017-0126).

Each Delphi study round consisted of two parts. *Part A* addressed the first research aim of developing consensus regarding the structure and definitions of the taxonomy. These questions were Likert scale responses (Strongly Disagree, Disagree, Neither Agree or Disagree, Agree and Strongly Agree). Participants who did not indicate agreement with an aspect were asked an open-ended question about what they would change with regards to the structure or definitions within the aspect. *Part B* explored the second research aim of examining application of the taxonomy by SLPs when describing assessments. Participants were asked to select taxonomy categories that they thought applied to assessments presented

in the Delphi study, with open-ended questions also provided for participants to make comments regarding the use of the taxonomy. In the first survey round, participants were asked to describe four assessments that were identified to them by name. These included: Clinical Evaluation of Language Fundamentals -4th Edition (CELF-4) (Wiig et al., 2004), Children's Communication Checklist – 2nd Edition (CCC-2) (Bishop, 2003), Language Sampling Protocol (Westerveld & Gillon, 2002), and the Pragmatics Profile of Everyday Communication Skills in Children (Dewart & Summers, 1995). Participants who identified themselves as being unfamiliar with any of the assessments were not required to provide categorisations for those assessments. At least 24 of the study participants categorised each assessment for each aspect of the taxonomy.

Analysis of data after round one led to the use of case studies in survey rounds two and three. The use of case studies made it possible for all participants to complete all questions as background information was provided about each assessment as well as a link to the published test website. The case studies were constructed to examine components of the taxonomy that may be more difficult to apply, while still being assessments that were characteristic of paediatric SLP practice. Two assessments were embedded within the case studies and the same case studies were used in both rounds two and three. Case study one described a parent interview using the Pragmatic Profile of Everyday Communication Skills for Children (Dewart & Summers, 1995) for a 4;10 year old child with Autism Spectrum Disorder. Case study two described a language sampling procedure using the Language Sampling Protocol (Westerveld & Gillon, 2002) with a 7;10 year old child experiencing language difficulties at school. These assessments were selected as they had the greatest inconsistency in agreement noted in round one compared to the agreement for the CELF-4 (Wiig et al., 2004) and CCC-2 (Bishop, 2003)

In each of three survey rounds, the proposed taxonomy structure and definitions were presented in a reference document along with a summary of background information and

references to literature. After each round, changes to the taxonomy in response to quantitative data (level of agreement with taxonomy structure and definitions or level of agreement with assessment categorisations) and qualitative data (themes from participant comments and feedback) were made by updating the taxonomy reference document. Changes were made with the aim of either increasing agreement with the taxonomy itself, or improving application of the taxonomy by addressing identified sources of confusion with definitions. In rounds two and three, participants were also provided with a document summarising the quantitative and qualitative group results from the previous round.

**3.3.3. Analysis**. Survey responses were analysed using the Statistical Package for the Social Sciences (SPSS version 20 software (IBM Corp, Released 2011). The number of rounds and criteria for agreed consensus were determined before the study commenced. In *Part A*, agreement with taxonomy structure and definitions was defined as 75% or more participants selecting "Strongly Agree" or "Agree" (i.e. median score of 4 or more on the five-point scale and inter-quartile range (IQR) of 1). In *Part B*, consensus on the categorisation of assessments was considered achieved when 75% or more participants selected (or opted not to select) a taxonomy category for an assessment. While agreement between the author's categorisations and Delphi participant's categorisations was not a requirement for consensus, examining concordance between the two provided an additional means of examining application of the taxonomy. Participant's survey responses to open-ended questions were analysed using conventional content analysis (Hsiu-Fang & Shannon, 2005) to identify themes for each aspect of the taxonomy. These themes were considered alongside quantitative data to identify potential reasons for lack of participant consensus (Tapio et al., 2011). Data analysis was conducted by the first author, who was blinded to the identity of participants during analysis, and results were reviewed by the other authors.

**3.4. Results**

    **3.4.1. Part A: Agreement with taxonomy structure and definitions**. Delphi participants reached consensus on the structure and definitions of the taxonomy in both rounds one and two, with 100% of participants strongly agreeing or agreeing with the overall structure of the taxonomy and at least 88.4% of participants strongly agreeing or agreeing with the sub-components and definitions within every aspect. No participants strongly disagreed with any aspect of the taxonomy. These results are provided in Table 3.2. As consensus was established across both rounds one and two, participants were not asked to rate their level of agreement regarding the structure and definitions in round three. Therefore, the round three survey only contained content related to *Part B*.

    **3.4.2. Part B: Categorisation of assessments using the taxonomy**. At the end of round three, consensus was established regarding seven out of the 10 components for case study one (parent interview for a child with Autism Spectrum Disorder) and four out of the 10 components for case study two (language sampling for a child experiencing language difficulties at school). The level of agreement with regards to the categorisation of case studies across each taxonomy component is provided in Table 3.3.

Table 3.2

*Participant Agreement with Structure of Taxonomy and Definitions (Part A)*

| Aspect of the taxonomy | Results | | | | | |
|---|---|---|---|---|---|---|
| | Median | | IQR | | Percentage agreement | |
| | Round 1 | Round 2 | Round 1 | Round 2 | Round 1 | Round 2 |
| | n=55 | n= 42 | n=55 | n=42 | n=55 | n=42 |
| Aspect I Structure | 4 | 5 | 1 | 1 | 92.7 (51) | 97.7 (41) |
| Aspect I Definitions | 4.5 | 4 | 1 | 1 | 90.9 (50) | 93.0 (39) |
| Aspect II Structure | 4 | 5 | 1 | 1 | 92.7 (51) | 95.2 (40) |
| Aspect II Definitions | 4 | 5 | 1 | 1 | 96.4 (53) | 93.0 (39) |
| Aspect III Structure | 4 | 5 | 1 | 1 | 87.3 (48) | 90.7 (38) |
| Aspect III Definitions | 4.5 | 5 | 1 | 1 | 90.9 (50) | 88.4 (37) |
| Aspect IV Structure | 4.5 | 5 | 1 | 1 | 96.4 (53) | 97.7 (41) |
| Aspect IV Definitions | 4.5 | 5 | 1 | 1 | 98.2 (54) | 95.2 (40) |
| [a] Overall Structure | 4 | 5 | 1 | 1 | 100 (55) | 100 (42) |

*Note:* Percentage agreement: Percentage of participants who selected "Agree" or "Strongly Agree"; Scale: 5 = Strongly Agree, 4 = Agree, 3 = Neither Agree or Disagree, 2 = Disagree, 1 = Strongly Disagree; Median: The value that appears most often (i.e., the most frequently selected answer); IQR: Inter-quartile Range i.e. the middle 50% of the data (i.e. the difference between 75th and 25th percentiles); [a] During round one, 54 participants completed this question.

Table 3.3

*Participant Agreement with Categorisation of Assessments in Case Studies (Part B)*

| Aspect within taxonomy | Categories within aspect | Results | | | |
|---|---|---|---|---|---|
| | | Case Study 1 | | Case Study 2 | |
| | | Round 2 n=43 % of participants who selected category | Round 3 n= 32 % of participants who selected category | Round 2 n=43 % of participants who selected category | Round 3 n=32 % of participants who selected category |
| **Aspect I** | Spoken | 97.7[a] | NA | 100[a] | NA |
| | Written | 2.3 | NA | 2.3 | NA |
| **Categories not mutually exclusive** | Semantics | **62.8[b]** | **37.5[b]** | **76.7[c]** | **56.3[b]** |
| | Morphosyntax | 7.0 | NA | 86.0[a] | NA |
| | Social Abilities | 100[a] | NA | **37.2** | NA |
| **In round three participants could only choose one category in addition the categories already agreed-upon in round two.** | Discourse | 18.6 | NA | 97.7[a] | NA |
| | Meta Abilities | 7.0 | NA | 18.6 | NA |
| | Executive Functions | **30.2[b]** | **28.1[b]** | **25.6[b]** | 25.0 |
| | Comprehension | 83.7[a] | NA | 100[a] | NA |
| | Production | 100[a] | NA | 100[a] | NA |
| **Aspect II** | Predict Outcome | **25.6[b]** | 25.0 | **58.1[ab]** | **34.4[ab]** |
| | Select Intervention | 79.1[a] | NA | **72.1[b]** | **43.8[b]** |
| **Categories not mutually exclusive.** | Plan Dosage | **39.5[b]** | 25.0 | **41.9[b]** | NA |
| **In round three participants could only choose one other prognostic and one other analytical category in addition to categories already agreed-upon in round two.** | Screening | **30.2[b]** | **31.3[b]** | 20.9 | NA |
| | Diagnostic | **41.9[b]** | **31.3[b]** | **46.5[b]** | 15.6 |
| | Detect Change | 23.3 | NA | **37.2[ab]** | 78.1[a] |
| | Describe Status | 87.7[a] | NA | 88.4[a] | NA |
| **Aspect III** | SLP Conducted | **39.5[b]** | 15.6 | 95.3[a] | NA |
| | Other Conducted | 0.0 | NA | 2.3 | NA |
| **Categories mutually exclusive** | Software | 0.0 | NA | 0.0 | NA |
| | Proxy – Reported | **60.[ab]** | 84.4[a] | 2.3 | NA |
| | Clinic | **34.9[b]** | **28.1[b]** | **53.5[ab]** | **53.1[ab]** |
| **Categories mutually exclusive** | Community – Home | **58.1[ab]** | **71.9[ab]** | 0.0 | NA |
| | Community – School | 0.0 | NA | **44.2[b]** | **46.9[b]** |
| | Community – Other | 7.0 | NA | 2.3 | NA |
| **Aspect IV** | Standardised | 20.9 | NA | **30.2[ab]** | **56.3[ab]** |
| **Categories mutually exclusive** | Non-standardised | 79.1[a] | NA | **69.8[b]** | **43.8[b]** |
| | Norm-referenced | 0.0 | NA | 7.0 | NA |
| **Categories mutually exclusive** | Criterion-referenced | 11.6 | NA | 14.0 | NA |
| | Descriptive data | 88.4[a] | NA | 79.1[a] | NA |
| | Static | 86.0[a] | NA | **39.5[b]** | **43.8[b]** |
| **Categories mutually exclusive** | Dynamic | 14.0 | NA | **60.5[ab]** | **56.3[ab]** |
| | Hierarchical | 9.3 | 9.4 | 4.7 | 15.6 |
| **Categories mutually exclusive** | Non-Hierarchical | 14.0 | 6.3 | 16.3 | 25.0 |
| | Contextualised | **65.1[ab]** | 78.1[a] | **48.8[ab]** | **56.3[ab]** |
| | Activity-focused | 11.6 | 6.3 | **30.2[b]** | 3.1 |

*Note.* Case Study 1: Parent interview using Pragmatics Profile; Case Study 2: Language sampling using Language Sampling Protocol; [a] Categories researchers expected would be selected for each case study; [b] Categories where inconsistency was identified i.e., 25-75% of participants selected this category; [c] Categories where inconsistency was identified due to an unexpected result i.e. this category reached criteria for consensus,

however consensus did not align with researcher expectations; NA: Not applicable as this question was not asked in round three due to consensus being reached in round two.

    **3.4.3. Final taxonomy.** The agreed-upon taxonomy has four aspects, labelled in roman numerals I-IV, each containing a number of components that describe assessments. The finalised structure of the taxonomy after round three is represented in Figure 3.1 and the finalised definitions of each taxonomy component after round three are provided in Supplementary Appendix 3.1. Each taxonomy aspect is described below followed by a summary of the components that were not consistently categorised in case studies at the end of round three. The themes identified from participant comments and associated changes to the taxonomy though the Delphi study rounds are summarised in Supplementary Appendix 3.2.
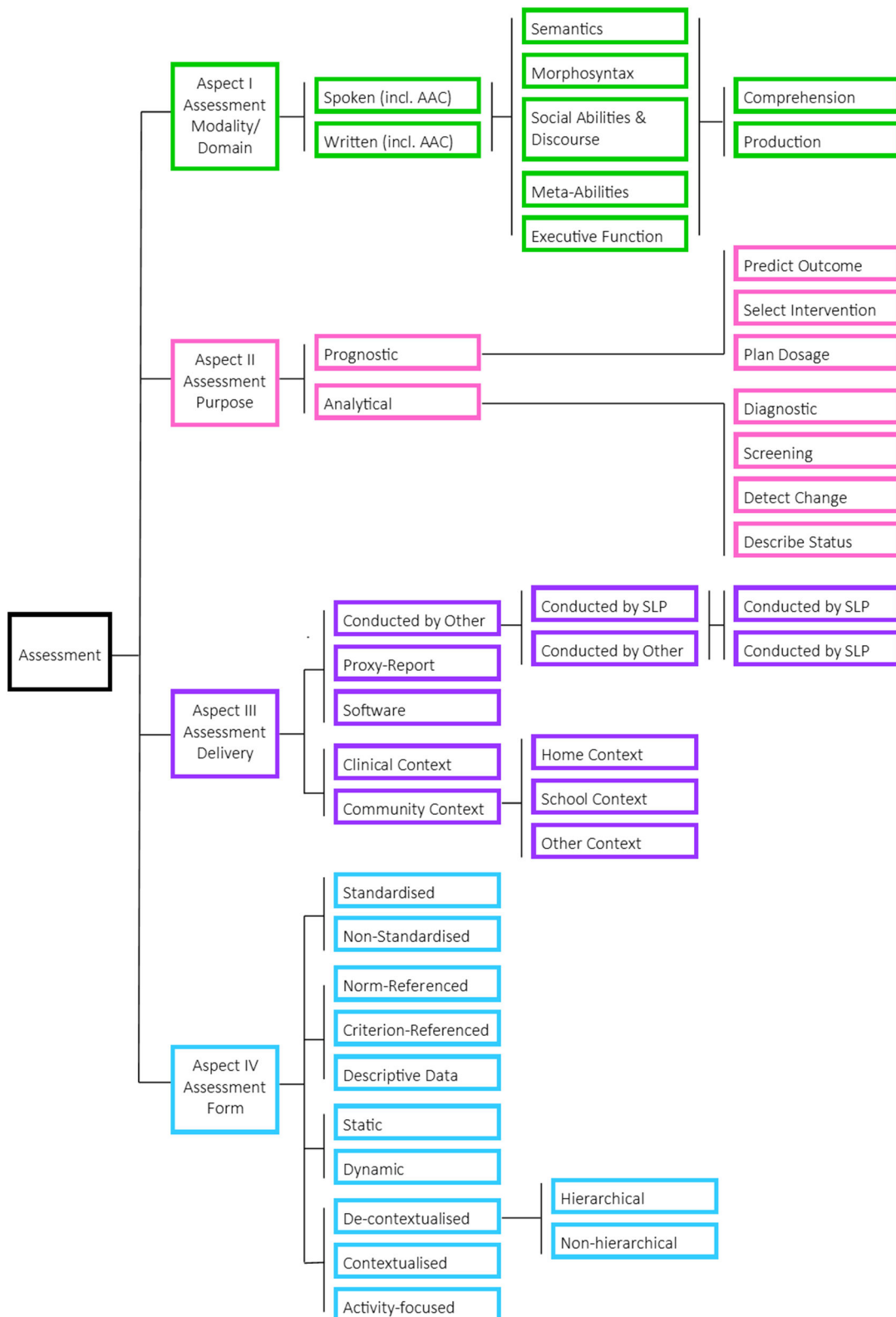
*Figure 3.1.* Taxonomy structure. *Note*: A different version of this same taxonomy is presented in Figure 4.1.

*3.4.3.1. Aspect I (Modalities and Domains).* Aspect I provides terminology for describing the skills that are specifically measured by an assessment and reported on in assessment findings. There are three components: modality, domain and comprehension/production. Modality includes categories spoken and written (including AAC). Domains include semantics, morphosyntax, social abilities & discourse, meta-abilities and executive functions. Assessments are also described as targeting comprehension (reception) or production (expression) of language. The categories in Aspect I are not mutually exclusive. An assessment may target either or both spoken and written modalities, either or both comprehension and production, and target one or more domains.

At the end of round three, the categories *semantics* and *executive functioning* remained inconsistently selected. Themes from participant comments suggested the following possible reasons for lack of consensus: (1) participants considering other ways an assessment could be conducted or selecting domains that may be involved in completing assessment tasks, but are not the key domains being measured by the assessment; (2) perceived overlap between the categories of *semantics* and *executive functioning*; and (3) the high level of information processing required from Delphi participants when reading and applying definitions.

*3.4.3.2. Aspect II (Assessment Purpose).* Aspect II describes the purposes for which assessments are used. There are seven categories that include predict outcome, select intervention, plan dosage (prognostic purposes relating to identification of possible future needs or supports) and screening, diagnostic, detect change and describe status (analytical purposes related to describing current functioning). These categories are not mutually exclusive as an assessment may have more than one purpose.

After round three, consensus on all Aspect II categories was not reached for either case study. Themes from participants comments identified the following reasons for lack of consensus: (1) the possibility of participants selecting all possible ways an assessment could

be used; (2) the potential for the *predict outcome* category to be only viewed as prognosis for improvement in diagnostic symptoms, rather than covering identification of future supports or need for intervention; and (3) individual SLP perceptions or service agency policy influencing categorisation. For example, if severity of diagnostic symptoms is used to determine eligibility for services within a particular clinical setting, then *diagnostic* purposes may not be differentiated from purposes of *predict outcome* or *select intervention*.

 ***3.4.3.3. Aspect III (Service Delivery).*** Aspect III provides terms for describing the methods and contexts in which assessments are conducted. This aspect has two components. The first component describes the method by which data is collected and includes three categories: (1) direct sampling, testing or observing a child's skills either by a SLP or by another trained person (e.g. teacher, parent or other professional), (2) assessment administered through a software program; and (3) collection of proxy-reported information (e.g. getting information from a parent through an interview or checklist. Assessments conducted by a SLP or a trained person can occur either face-face or via telehealth using information and communication technologies (ICTs). The other component in Aspect III considers the environmental context targeted in the assessment. Assessment may occur within a clinical context or within home, school or other community contexts (Parsons et al., 2005; Schraeder, Quinn, Stockman, & Miller, 1999). Each category in Aspect III is mutually exclusive from other categories within each of the respective components (i.e. a particular assessment is conducted via only one method and targets only one environmental context).

 After round three, lack of consensus remained with regards to the *environmental context* categories for both case studies. Participant comments identified: (1) lack of clarity between the *environmental context* targeted in the assessment and the physical location of the assessment; (2) possible confusion between *environmental context* and the *task type* categories in Aspect IV (Form); (3) participants focussing on one element of an assessment

rather than selecting the category that best fits overall; and (4) the high level of information processing required by Delphi participants when applying definitions to case studies.

*3.4.3.4. Aspect IV (Form).* Aspect IV has four components. These include: (1) a component that distinguishes between standardised or non-standardised administration procedures; (2) a component that describes the type of data collected in terms of norm-referenced, criterion referenced or descriptive; (3) a component that identifies assessments as either static or dynamic; and (4) a component that describes task type in terms of discrete skill tasks versus contextualised or performance-based tasks and the naturalness of the communication interaction during assessment tasks. Each category in Aspect IV is mutually exclusive from other categories within each of the respective components (i.e. a particular assessment is either standardised or non-standardised; either norm-referenced, criterion-referenced or descriptive; either static or dynamic and is one task type).

At the end of round three, consensus on case study one was achieved with regards to all Aspect IV components, however case study two lacked consensus. Participants comments reflected the following explanations for lack of consensus: (1) SLPs selected all possible ways an assessment could be used, rather than considering only the purposes for which assessments were used in cases studies; (2) it may be difficult to distinguish between the *task type* categories, *contextualised* and *activity-focussed*; (3) participants may ascribe definitions that are different to the taxonomy definitions when applying assessments in the respective case studies; and (4) the high level of information processing required from Delphi participants when applying the taxonomy definitions to case studies.

## 3.5. Discussion

In this study, a taxonomy for describing language assessments was developed, with experienced SLPs from a variety of work sectors reaching consensus on categories and definitions for describing the key features of assessments. Given the numerous challenges associated with the development of agreed-upon terminology, including the wide array of

activities that may be undertaken when assessing the language abilities of children and the varied work sectors that span paediatric SLP practice (Walsh, 2005); this taxonomy represents a significant advancement in the field of child language assessment.

The use of case studies in the study allowed the application of the taxonomy to be examined and, in doing so, facilitated the refinement of the terms and definitions within the taxonomy. Nonetheless consensus across all components of the taxonomy with regards to categorisation of assessments was not reached for either case study. Greater inconsistency existed for case study two, particularly with regards to Aspect IV. Case study two described a language sampling procedure that followed a standardised procedure, but was dynamic in nature and provided descriptive data. It was noted in round one that assessment tools that are less prescriptive and more variable in terms of how they might be used, were less likely to be categorised consistently. The resource used in case study two was the Language Sampling Protocol (Westerveld & Gillon, 2002), a tool which may potentially have wide and varied applications by SLPs. It is possible that, despite all having the same case study, participants were drawn to considering how they themselves use the assessment tool, rather than how the tool's use was described in the case study. This may have contributed to this case study being less consistently categorised.

The components of the taxonomy that were not categorised consistently in case studies may also represent areas of SLP theoretical understanding that may need further development within the profession. For example, a lack of agreement on whether the assessments in the case studies targeted *semantics* and to a lesser extent, *executive functioning* persisted after round three. This lack of agreement may go beyond terminology and could reflect differences in professional understanding with regards to how these domains are assessed.

*Environmental context* also lacked consensus in both case studies, despite attempts to clarify this across Delphi rounds. While SLP literature discusses the value of assessing skills

in everyday communication environments, this discussion often occurs in the context of specific examples using terminology such as authentic or curriculum-based (Parsons et al., 2005; Schraeder et al., 1999). Similarly, while the concept of dynamic assessment is discussed across literature (Dockrell & Marshall, 2015), it is often presented as an assessment approach for children learning English as a second language and thus may not be an approach that SLPs in general paediatric language practice frequently identify themselves as using (Caesar & Kohler, 2009). This creates the possibility that, while SLPs agreed with the taxonomy distinctions for *environmental context* and *dynamic* assessment; applying the taxonomy may require SLPs to make more explicit and specific distinctions between assessments than they have previously been accustomed to making.

The identification of purposes for which assessments are used also emerged as an area of inconsistency in case studies, with participants tending to select many purposes for a single assessment. While it is important that assessment data be used maximally, it is also important that assessments are used for the purposes for which they have been designed (Newton, 2007). Researchers and clinicians must carefully decide which psychometric properties are most essential for a particular purpose and, thus, are most important to focus upon when selecting an assessment for that purpose (Wade, 2004). For example, assessments used for *diagnostic* and *screening* purposes should have established sensitivity and specificity data, whilst assessments used for *detecting change* should have evidence of responsiveness (Wade, 2004). While the extant SLP literature has focussed on assessments suited for *diagnostic* and *screening* purposes (Dockrell & Marshall, 2015), there appears to be limited literature explicitly examining assessments for purposes other than these, with most of the literature on assessment purpose originating from literature outside the SLP discipline (Newton, 2007; Wade, 2004). In the future, greater attention may need to be places on the purposes of different language assessments for SLP professional knowledge of language assessment practice to develop. Limitations with regards to current availability of assessments with

established psychometric properties, particularly assessments that target communication from a more functional perspective (McLeod & Threats, 2008); may also lead to the use of assessments that are not ideally suited to the clinical purpose. It is also possible that constraints such as the high cost of commercial assessments, limited time to conduct assessments, or service provider policy demands may also overshadow decisions regarding the purposes for which assessment data is to be used (Fulcher-Rood, Castilla-Earls, & Higginbotham, 2018). Further investigation of assessment practices is necessary, particularly with regards to factors that influence SLP choice of assessment.

The findings of this Delphi study show that, even when consensus was reached on the categories and definitions within the taxonomy, consistent application of terminology cannot be assumed. The field of implementation science acknowledges challenges with knowledge to action transfer and the successful adoption of practice innovations (Miao, Power, & O'Halloran, 2015; Wilson, Brady, Lesesne, & NCCDPHP Work Group on Translation, 2011). These same challenges may apply to the adoption of new terminology. Although use of the taxonomy does not involve change to clinical practice per se, it may require SLPs to use new terminology or define terms related to assessment differently to what they may be accustomed to. Some terms may be engrained in particular organisations, service providers or in the minds of individual SLPs. In those circumstances, SLPs may need to develop an explicit understanding of how terminology in the taxonomy relates to the terminology they currently use in order to effectively 'code switch' between terms. With this in mind, further research is needed to identify specific actions and strategies to assist consistent application of the taxonomy by SLPs when describing clinical practice (Wilson et al., 2011).

**3.5.1. Limitations**. Participants in this study represented a variety of geographical locations, work sectors and levels of professional experience, however, as with any Delphi study, it cannot be assumed that the same findings would be reached with a different group of participants. This study was also limited to Australia. Given that almost all of the background

literature and research associated with the taxonomy originates from the United States or United Kingdom, it is expected that the terms in the taxonomy would also be applicable to other English-speaking countries; however further research is warranted.

Participant drop-out over rounds poses a limitation in Delphi Studies (Boulkedid et al., 2011). In this study, agreement with the structure and definitions of the taxonomy was reached in round one, with 55 participants. Completion rates for round two and three were 74.4% and 78.2% respectively. Whilst this completion rate is reported as being typical in web-based surveys (Schleyer & Forrest, 2000), it may pose a limitation for *Part B* of the study in which categorisation of assessments using the taxonomy was examined; as it cannot be presumed that drop-out was random. Further, there was a large amount of reading required from participants in completing the surveys in this study, particularly in *Part B* which required reading the survey questions and taxonomy definitions alongside the case studies. While all attempts were made to present information in a reader-friendly manner, it is possible that categorisation of case studies was influenced by participant fatigue associated with high cognitive demand.

While the use of case studies served a purpose of allowing application of the taxonomy to be examined, the case studies are not without limitations themselves. It is possible that the use of case studies may have drawn participants to considering the case (i.e. describing the domains that may require assessment based on the child's needs), rather than describing the specific assessment used in the case study. It was also not possible to comprehensively examine all aspects of the taxonomy using two case studies. Results may have been different if the case studies used other types of assessments.

### 3.6. Conclusion

In this study, a taxonomy for describing child language assessment practices was developed and a rigorous methodology applied in order to evaluate the consensus of it amongst a group of experienced paediatric SLPs. The high level of agreement from clinicians

and academics with the taxonomy structure and definitions represents a significant step towards addressing some of the challenges that inconsistent terminology poses for the field of child language assessment. The taxonomy provides structure, terminology and definitions from which further professional knowledge and future research may be built upon (Eadie, 2003). It has uses for the collection of data on SLP assessment practices, provision of SLP training, and for making comparisons between different assessments in research studies. Given that some components of the taxonomy were not consistently applied when describing the case studies, further research is recommended to identify strategies that support implementation of the taxonomy.

**References for Chapter 3**

American Speech and Hearing Association. (2018). *Assessment tools, techniques, and data sources* Retrieved from https://www.asha.org/Practice-Portal/Clinical-Topics/Late-Language-Emergence/Assessment-Tools-Techniques-and-Data-Sources/

Barnes, S., & Bloch, S. (2018). Why is measuring communication difficult? A critical review of current speech pathology concepts and measures. *Clinical Linguistics & Phonetics, 33*(3), 219-236. doi:https://dx.doi.org/10.1080/02699206.2018.1498541

Bishop, D. V. M. (2003). *Children's Communication Checklist - Second Edition*. London, England: Psycholgical Corporation.

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., & Alberti, C. (2011). Using and reporting the Delphi method for selecting healthcare quality indicators: A systematic review. *PLoS ONE, 6*(6), e20476. doi:10.1371/journal.pone.0020476

Caesar, L. G., & Kohler, P. D. (2009). Tools clinicians use: A survey of language assessment procedures used by school-based speech-pathologists. *Communication Disorders Quarterly, 30*(4), 226-236. doi:10.1177/1525740108326334

Cowie, M., Wanger, K. M., Cartwright, A., Bailey, H., Millar, J. A., & Price, M. (2001). A review of Clinical Terms Version 3 (Read Codes) for speech and language record keeping. *International Journal of Language & Communication Disorders, 36*(1), 117-126. doi:10.1080/13682820116848

DeJong, G., Horn, S. D., Gassaway, J. A., Slavin, M. D., & Dijkers, M. P. (2004). Toward a taxonomy of rehabilitation interventions: Using an inductive approach to examine the "black box" of rehabilitation. *Archives of Physical Medicine and Rehabilitation, 85*(4), 678-686. doi:10.1016/j.apmr.2003.06.033

Dewart, H., & Summers, S. (1995). Pragmatics Profile of Everyday Communication Skills in Children. Retrieved from http://complexneeds.org.uk/modules/Module-2.4-

Assessment-monitoring-and-

evaluation/All/downloads/m08p080c/the_pragmatics_profile.pdf

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in

young children. *Child and Adolescent Mental Health, 20*(2), 116-125.

doi:10.1111/camh.12072

Eadie, P. (2003). Speech pathology assessment practices: One assessment or many? *Advances

in Speech Language Pathology, 5*(1), 65-68. doi:10.1080/14417040510001669081

Fulcher-Rood, K., Castilla-Earls, A. P., & Higginbotham, J. (2018). School-based speech-

language pathologists' perspectives on diagnostic decision making. *American Journal

of Speech-Language Pathology, 27*(2), 796-812. doi:10.1044/2018_AJSLP-16-0121

Hegde, M. N., & Pomaville, F. (2017). Assessment of communication disorders in children.

In *Assessment of communication disorders in children: Resources and protocols* (pp.

3-16). Sa Diego, CA: Plural Publishing.

Hsiu-Fang, H., & Shannon, S. E. (2005). Three approaches to qualitative content analysis.

*Qualitative Health Research, 16*(9), 1277-1288. doi:10.1177/1049732305276687

Hughes, J., & Orange, J. B. (2007). Mapping functional communication measurements for

traumatic brain injury to the WHO-ICF. *Revue canadienne d'orthophonie et

d'audiologie, 31*(3), 135-142.

IBM Corp. (Released 2011). IBM SPSS Statistics for Windows Version 20.0. Armonk, NY:

IBM Corp.

Kaderavek, J. N. (2015). Assessment of language disorders. In *Language disorders in

children: Fundamental concepts of assessment and intervention* (pp. 45-95). US:

Pearson.

Lahey, M. (1988). *Language disorders and language development*. New York: MacMillan.

Larson, V. L., & McKinley, N. L. (2003). Procedures for direct assessment. In *Communication solutions for older students* (pp. 209-268). Eau Claire, WI: Thinking Publications.

McLeod, S., & Threats, T. T. (2008). The ICF-CY and children with communication disabilities. *International Journal of Speech-Language Pathology, 10*, 92–109. doi:10.1080/1754950070183469

Miao, M., Power, E., & O'Halloran, R. (2015). Factors affecting speech pathologists' implementation of stroke management guidelines: A thematic analysis. *Disability and Rehabilitation, 37*(8), 674-685. doi:10.3109/09638288.2014.932444

Newton, P. E. (2007). Clarifying the purposes of educational assessment *Assessment in Education: Principles, Policy & Practice, 14*(2), 149-170. doi:10.1080/09695940701478321

Parsons, S., Law, J., & Gascoigne, M. (2005). Teaching receptive vocabulary to children with Specific Ianguage Impairment: A curriculum-based approach. *Child Language Teaching and Therapy, 21*(1), 39-59. doi:10.1191/0265659005ct280oa

Paul, R., & Norbury, C. F. (2012). Chapter 2: Assessment. In R. Paul & C. F. Norbury (Eds.), *Language disorders from infancy through adolescence: Listening, speaking, reading, writing and communicating* (4th ed., pp. 22-60). St Louis, MI: Mosby Elsevier.

Roulstone, S. (2001). Consensus and variation between speech and language therapists in the assessment and selection of preschool children for intervention: A body of knowledge or idiosyncratic decisions? *International Journal of Language and Communication Disorders, 36*(3), 329–348. doi:10.1080/13682820010019928

Roulstone, S. (2015). Exploring the relationship between client perspectives, clinical expertise and research evidence. *International Journal of SpeechLanguage Pathology, 17*(3), 211-222. doi:10.3109/17549507.2015.1016112

Schleyer, T. K., & Forrest, J. L. (2000). Methods for the design and administration of web-based surveys. *Journal of the American Medical Informatics Association, 7*(4), 416-425. doi:10.1136/jamia.2000.0070416

Schraeder, T., Quinn, M., Stockman, I. J., & Miller, J. (1999). Authentic assessment as an approach to preschool speech-language screening. *American Journal of Speech-Language Pathology, 8*(3), 195-200. doi:10.1044/1058-0360.0803.195

Tapio, P., Paloniemi, R., Varho, V., & Vinnari, M. (2011). The unholy marriage? Integrating qualitative and quantitative information in Delphi processes *Technological Forecasting & Social Change, 78*, 1616 −1628. doi:10.1016/j.techfore.2011.03.016

Wade, D. T. (2004). Assessment, measurement and data collection tools. *Clinical Rehabilitation, 18*, 233-237. doi:10.1191/0269215504cr183ed

Walsh, R. (2005). Meaning and purpose: A conceptual model for speech pathology terminology. *Advances in Speech-Language Pathology, 7*(2), 65-76. doi:10.1080/14417040500125285

Walsh, R., & IGOTF-CSD. (2006). A history of terminology: International group on terminology frameworks – communication science and disorders. . Retrieved from http://www.dhrs.uct.ac.za/sites/default/files/image_tool/images/147/History_of_CSD.pdf

Westby, C., & Washington, K. N. (2017). Using the International Classification of Functioning, Disability and Health in assessment and intervention of school-aged children with language impairments. *Language, Speech, and Hearing Services in Schools, 48*, 137-152. doi:10.1044/2017_LSHSS-16-0037

Westerveld, M. F., & Claessen, M. (2014). Clinician survey of language sampling practices in Australia. *International Journal of Speech-Language Pathology, 16*(3), 242-249. doi:10.3109/17549507.2013.871336

Westerveld, M. F., & Gillon, G. T. (2002). Language Sampling Protocol. Retrieved from http://www.education.canterbury.ac.nz/documents/gillon/languageprotocol.pdf

Wiig, E. H., Semel, E., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals - 4th Edition* (5 ed.). Bloomington, MN: Pearson.

Wilson, K. M., Brady, T. J., Lesesne, C., & NCCDPHP Work Group on Translation. (2011). An organizing framework for translation in public health: The knowledge to action framework. *Preventing Chronic Disease, 8*(2), 1-7.

World Health Organisation. (2015). *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (Vol. 10). Geneva, Switzerland: World Health Organisation.

**Supplementary Appendix 3.1.**

Taxonomy Categories and Definitions

<table>
<tr>
<td colspan="2" align="center"><b>ASPECT I</b><br><b>(Modalities and Domains)</b></td>
</tr>
<tr>
<td align="center"><b>Term and definition</b></td>
<td align="center"><b>Examples of assessments</b></td>
</tr>
<tr>
<td><b>Spoken Language:</b><br>Language exchanged verbally, or via an alternative in situations where peers would typically use verbal communication (includes pre-linguistic communication)<br><br>(American Speech and Hearing Association, 1993; Beukelman & Mirenda, 2013).</td>
<td>• <i>Assessment of spoken communication via a single mode (single-modality) e.g. Speech-only or AUSLAN</i><br>• <i>Assessment of spoken communication via multiple modes (multi-modal) e.g. Key-word sign or Aided language stimulation</i></td>
</tr>
<tr>
<td><b>Written Language:</b><br>Language exchanged through text (print) or via an alternative in situations where peers would be typically be reading or writing<br><br>(American Speech and Hearing Association, 1993; Beukelman & Mirenda, 2013).</td>
<td>• <i>Assessment of written communication via a single mode (single-modality) e.g. Text-only</i><br>• <i>Assessment of written communication via multiple modes (multi-modal) e.g. Text with symbol support</i></td>
</tr>
<tr>
<td><b>Semantics:</b><br>Understanding and expression of words and word meanings (e.g. vocabulary, word retrieval, lexical meaning).<br><br>(American Speech and Hearing Association, 1993; Apel, 2014; Beukelman & Mirenda, 2013; Dockrell & Marshall, 2015; Paul & Norbury, 2012).</td>
<td>• <i>Knowledge of vocabulary words is assessed by having the child name a series of pictures</i><br>• <i>A sample of a child's language is analysed for number of different words (NWD) or type-token ratio (TTR)</i><br>• <i>Semantic knowledge is assessed by asking the child to give synonyms and antonyms for different words</i></td>
</tr>
<tr>
<td><b>Morphosyntax:</b><br>Understanding and expression of different word forms and the order and combination of words in sentences<br><br>(American Speech and Hearing Association, 1993; Apel, 2014; Beukelman & Mirenda, 2013; Dockrell & Marshall, 2015; Paul & Norbury, 2012).</td>
<td>• <i>Sentence structure is assessed by asking a child to point to pictures that represent a spoken sentence</i><br>• <i>A sample of a child's language is analysed for MLU and Brown's Grammatical Morphemes</i></td>
</tr>
</table>

| | |
|---|---|
| **Social Abilities and Discourse (Pragmatics)**: Giving and making meaning in social context or communication for social purposes.  Includes:<br>- **Pre-linguistic communication** *e.g. facial expression, joint attention, gesturing etc*<br>- **Communication intentions/purposes** *e.g. Requesting, commenting, greetings, asking questions, giving reasons, making predictions etc*<br>- **Non-verbal communication** *e.g. understanding emotions from body language and facial expressions*<br>- **Non-literal language** *e.g. inferences, idioms, metaphors, jokes, sarcasm etc*<br>- **Matching communication style to social context** *e.g. Adjusting communication style between friends and teachers*<br>- **Conversation conventions** *e.g. topic selection, topic maintenance, conversational turn-taking etc*<br>- **Text cohesion** *e.g. verbal fluency (mazes and incomplete sentences), transitions between sentences/paragraphs etc*<br>- **Text organisation** (discourse or macrostructure) *e.g. Narrative structure (story grammar), episodic structure etc*<br><br>(American Speech and Hearing Association, 1993; Apel, 2014; Beukelman & Mirenda, 2013; Dockrell & Marshall, 2015; Paul & Norbury, 2012). | • *Information on the range of communication functions for which a child communicates is profiled during a parent interview*<br>• *Narrative structure (story grammar) and text cohesion are assessed during a narrative retell task*<br>• *Non-verbal communication and conversation conventions are observed during a conversation between the child and the SLP* |
| **Meta-Abilities:** Ability to think about own thought processes and understand how to regulate these processes for effective learning. Includes:<br>- **Meta-cognition**: Knowledge and use of strategies for managing and self-monitoring own learning.<br>- **Meta-Language:** Knowledge of phonemic (phonemic awareness), morphological/syntactic (meta-syntactic) or text-level (meta-narrative) rules in relation to own skills; and ability to effectively apply these rules for improved performance.<br>- **Meta-pragmatics:** Knowledge of social conventions in relation to own communication and ability to apply this knowledge to improve communication with others<br><br>(Kamhi, Masterson, & Apel, 2007; Larson & McKinley, 2003; Law, Campbell, Roulstone, Adams, & Boyle, 2008; Starling, Munro, Togher, & Arciuli, 2012) | • *A child is asked to describe strategies that facilitate their own learning or performance (meta-cognition)*<br>• *A child describes the features of a narrative story and their understanding of what constitutes good narrative structure (meta-language)*<br>• *Phoneme awareness skills are assessed by asking the child to identity the number of phonemes in words (meta-language)*<br>• *A child is asked to identify what they would do in a given social situation and why (meta-pragmatics)* |

| | |
|---|---|
| **Executive Functions:**<br>Collection of related cognitive processes necessary for execution of goal-directed, controlled, purposeful behaviour. Includes:<br>- **Inhibition** (self-control): Ability to focus and attend to tasks through suppression of inappropriate thoughts, comments and behaviours<br>- **Emotion control** (self-regulation): Ability to manage emotions for goal achievement and task completion<br>- **Working memory**: Ability to retain, process and manipulate pieces of information for short periods of time to complete required tasks<br>- **Organisation:** (strategic planning) Ability to use organisational strategies for task completion e.g. envisioning the end product, planning steps to complete tasks, identifying solutions to problems etc<br>- **Mental flexibility**: Ability to integrate prior knowledge and experiences or effectively apply of different rules for different situations<br>- **Sustained attention**: Ability to maintain attention to tasks despite distractions or fatigue<br><br>(Hyter, 2003; Montgomery, Magimairaj, & Finney, 2010; Serry, Rose, & Liamputtong, 2008; Singer & Bashir, 1999; Ukrainetz, 2006) | • *Auditory working memory is assessed by asking the child to repeat strings of numbers or words*<br>• *Organisational skills are assessed by observing a child in class while they plan out a project by setting goals and identifying steps involved.*<br>• *Inhibitory control is examined through a task that requires the child to read names of colours written in coloured ink that does not match the word that's spelled out i.e. the child must say the colour they see, as opposed to the word that is written* |
| **Comprehension:** Understanding of information, knowledge and ideas communicated by others (includes verbal and non-verbal).<br><br>(American Speech and Hearing Association, 1993) | • *A child's ability to understanding and follow directions is assessed by asking the child to follow a series of instructions*<br>• *A child's understanding of facial expressions is assessed by asking the child to point to faces that display different emotions* |
| **Production:** Ability to convey information, knowledge and ideas to others (includes verbal or non-verbal).<br><br>(American Speech and Hearing Association, 1993) | • *A child's vocabulary is assessed in a picture naming task*<br>• *A child's ability to produce a story is assessed in a narrative retell task* |

|  |
|:---:|
| **ASPECT II**<br>(Assessment Purpose) |

| **Term and definition** | **Examples** |
|---|---|
| **Predict outcome:**<br>Identify risk of poor future outcome, predict need for intervention or identify support needs.<br><br>(Olswang & Bain, 1996; Vaz et al., 2015; Wade, 2004; Wixson & Valencia, 2011) | • *Support needs at school (type/level of curriculum differentiation or special education support) are identified by assessing performance in the presence of different prompts or scaffolds (i.e. dynamic assessment using graded prompting).*<br>• *Early primary school or kindergarten children are assessed on pre-literacy skills that are seen as predictive of later literacy success (to identify those who may benefit from participation in a preventive program)* |

| | |
|---|---|
| **Select intervention:**<br>Identify suitability for an intervention approach or select intervention targets.<br><br>(Eadie, 2003; Newton, 2007; Paul & Norbury, 2012; Vaz et al., 2015; Wade, 2004; Westerveld & Claessen, 2014; Wixson & Valencia, 2011) | • *An interview with parents (regarding family preferences/concerns, child's likes/dislikes, available resources etc) assists with selection an intervention approach.*<br>• *A child's ability to produce a range of different morphological and syntactical forms is assessed to identify the forms to be targeted in intervention.* |
| **Plan dosage:** Predict intensity (dosage) of intervention.<br><br>(Vaz et al., 2015; Wade, 2004; Westby, 2007) | *The amount of intervention needed to achieve an outcome is estimated by:*<br>• *Assessing a child's response to a short trial of the intervention (dynamic assessment in a test-teach-retest format)*<br>• *Collecting a comprehensive history regarding the child's response to previous interventions (response to intervention).* |
| **Screening:**<br>Identify children who may have a disorder that requires further diagnostic assessment to confirm.<br><br>(American Speech-Language-Hearing Association, 2004; Dockrell & Marshall, 2015; Eadie, 2003; Paul & Norbury, 2012; Vaz et al., 2015; Wade, 2004; Westerveld & Claessen, 2014; Wixson & Valencia, 2011) | • *Assessment is conducted to identify if diagnostic assessment should be conducted and/or the domains to be targeted in diagnostic assessment* |
| **Diagnostic:**<br>Diagnose a condition or make a comparison with peers.<br><br>(American Speech-Language-Hearing Association, 2004; Betz, Eickhoff, & Sullivan, 2013; Dockrell & Marshall, 2015; Eadie, 2003; Kapantzoglou, Restrepo, & Thompson, 2012; Paul & Norbury, 2012; Vaz et al., 2015; Wade, 2004; Westerveld & Claessen, 2014; Wixson & Valencia, 2011) | • *Assessment conducted to identify the presence or severity of a diagnosis; or determine if functioning is different to peers* |
| **Detect change:**<br>Measure change in status or monitor progress over time.<br><br>(Eadie, 2003; Paul & Norbury, 2012; Vaz et al., 2015; Wade, 2004; Westerveld & Claessen, 2014; Wixson & Valencia, 2011) | • *Assessment repeated at different intervals to monitor progress over time*<br>• *Pre & post intervention assessment to document change (or no change in a control group)* |
| **Describe status:**<br>Assessment for the purpose of describing or explaining a particular aspect of a student's functioning.<br><br>(Vaz et al., 2015; Wade, 2004; Wixson & Valencia, 2011) | • *Communicative behaviours are described (gesture dictionary) in order to help unfamiliar communication partners understand/interpret a student's communication behaviours*<br>• *An SLP assesses a student's performance on spoken comprehension tasks to further explore reasons why others report that the student has difficulties understanding verbal information, despite the student achieving an average score on a standardised receptive language test.* |

| **ASPECT III** <br> **(Assessment Delivery)** | |
|---|---|
| **Term and definition** | **Examples** |
| **By Person - Conducted by SLP:** <br> Assessment is conducted by an SLP through pre-planned observation, testing or sampling of a child's skills. Results may be analysed at the time or may be analysed later from an audio/video recording. Others may assist with administration or technology may be used to score; however, the SLP has the primary role in planning the assessment and analysing findings. <br><br> (Kaminski, Abbott, Aguayo, Latimer, & Good, 2014; Wixson & Valencia, 2011) | • *An SLP conducts a standardised assessment* <br> • *An SLP transcribes and analyses a language sample that was audio-recorded earlier by a teacher* <br> • *An SLP compares and analyses a narrative transcript with reference to a database of normative data from peers* |
| **By Person - Conducted by Other:** Assessment conducted by another person (teacher, another professional etc), through pre-planned observation, testing or sampling of the child's skills. An SLP may provide training or support to the other person, or technology may be used (e.g. online stimulus materials or software that calculates test scores); however, the other person has the primary role in planning the assessment and analysing/interpreting results. <br><br> (Kaminski et al., 2014; Wixson & Valencia, 2011) | • *A teacher assesses the phonemic awareness skills of a group of children with literacy difficulties to determine literacy intervention goals for those children.* |
| **Face-to-face (only for assessments conducted by a person):** Assessment is conducted with the child and an assessor in the same room. <br><br> (Edwards, Stredler-Brown, & Houston, 2012; Mashima & Doarn, 2009; Waite, Theodoros, Russell, & Cahill, 2010a, 2010b) | • *During a face-to-face interaction with a child, an SLP audio-records a language sample for later analysis* <br> • *An SLP administers a standardised test face-to-face and scores with the assistance of scoring software* |
| **ICT (only for assessments conducted by a person):** Assessments is conducted with the assessor and the child communicating through ICTs (information and communication technologies). Technology that is not used for two-way communication between individuals during the assessment is not considered ICT (e.g. audio/video recorders) <br><br> (Edwards et al., 2012; Mashima & Doarn, 2009; Waite et al., 2010a, 2010b) | *Assessments conducted by:* <br> • *Web-conferencing (such as Skype or Zoom)* <br> • *Video-conferencing* <br> • *Telephone* |

| | |
|---|---|
| **Proxy-Report:** Skills are not assessed in the moment they occur, but are documented based on retrospective reports from others, such as in an interview or by completion of questionnaire/checklist. The reported information:<br>- may be from a child (self-report), another professional, a caregiver, a teacher or a peer.<br>- may relate to previous skills (e.g. developmental or educational history) or current abilities (e.g. current level of development; or performance in the current unit of schoolwork).<br><br>(Bishop & McDonald, 2009; Dockrell & Marshall, 2015; Schraeder, 2008; Williams, 2006) | • *During a case history interview, a parent reports on information about a child's history that may be diagnostically significant.*<br>• *A teacher reports information by completing a checklist regarding the pre-linguistic behaviours they have observed the child use at school.* |
| **Software delivered:** The child's abilities are assessed through a predominantly computerised procedure with no (or extremely little) input from a person. Software program selects/presents tasks, records data and scores results. A person may set a child up at a computer or be present to supervise while the child sits the test. If a person is required to administer items, respond to the child's test answers, record observations or score results; then the assessment is not categorised as software.<br><br>(Ockey, 2009; Richards et al., 2017) | *Assessments conducted by:*<br>• *App or web-based program*<br>• *Computer (software) program* |
| **Clinical context:** Skills are assessed within a clinical context i.e. the assessment does not incorporate materials or communication partners from the day-to-day environment.<br>Note: This category refers to the context being assessed; not physical location<br><br>(Schraeder, Quinn, Stockman, & Miller, 1999; Westby, 2007) | • *A child is withdrawn from regular classroom activities for narrative assessment by an SLP using materials that the SLP has brought to the school. Although the child is at school, the assessment context is that of a clinical environment*<br>• *An SLP administers a standardised assessment at the child's home in a quiet room away from distractions. The assessment is conducted according to administration guidelines and does not incorporate any of the activities, materials or people that the child interacts with at home.* |
| **School context:**<br>Communication is assessed in a school (or Kindergarten) context i.e. uses communication partners, communication situations or materials that represent a school environment.<br>Note: This category refers to the context being assessed; not physical location<br><br>(Nelson, 1989; Parsons, Law, & Gascoigne, 2005; Schraeder, 2008; Schraeder et al., 1999; Westby, 2007) | • *In an interview with the SLP, a teacher is asked to comment on how the child communicates with teachers and classmates during whole class lessons*<br>• *An SLP assesses a child's oral and reading comprehension skills using the text being studied in the current unit of English and activities similar to those used to teach the English school curriculum* |
| **Home context:**<br>Communication is assessed in a home context i.e. uses communication partners, communication situations or materials that represent a home environment.<br>Note: This category refers to the context being assessed; not physical location<br><br>(Schraeder, 2008; Schraeder et al., 1999; Westby, 2007) | • *During an appointment in an outpatient clinic, a parent completes a checklist based on the communication behaviours they have observed at home*<br>• *An SLP observes a child play and read with his mother and siblings using similar toys and books as those in the child's home. Although the child is in a clinic consultation room, the assessment context is considered to be representative of a home environment* |

| **Other community context:** Communication is assessed in a community context i.e. uses communication partners, communication situations or materials that represent a community environment. <br> Note: This category refers to the context being assessed; not physical location <br><br> (Schraeder, 2008; Schraeder et al., 1999; Westby, 2007) | • *A child describes the communication difficulties they experience when interacting coaches and teammates during extra-curricular soccer training and matches* <br> • *A child's skills are observed and recorded during a work experience placement (e.g. interacting with customers, taking orders and counting money)* |
|---|---|

<div align="center">

**ASPECT IV**
**(Assessment Form)**

</div>

| **Term and definition** | **Examples** |
|---|---|
| **Standardised:** <br> Assessments designed to be administered and scored in a consistent manner, which is the same for all children who are assessed i.e. specific questions or tasks, clear administration and scoring guidelines, defined assessment materials and set procedures to elicit responses from the child. <br><br> (Hegde & Pomaville, 2017; Paul & Norbury, 2012) | • *Use of a language sampling that follows specific administration procedures, including use of set materials and specific prompts to elicit the retell from the child* |
| **Non-standardised:** <br> Assessments that may not be administered the same way by different assessors in different conditions. Procedures for administration and scoring may be variable or may not be described well enough for consistent administration and scoring. <br><br> (Hegde & Pomaville, 2017; Paul & Norbury, 2012) | • *Use of a language sampling procedure that does not have set administration guidelines i.e. a task that the SLP has created themselves or adapted from another resource.* |
| **Norm-referenced:** <br> Assessments that quantitatively compare a child's score to scores from a sample of matched peers who completed the same task. These assessments should always be standardised. <br><br> (Caesar & Kohler, 2009; Paul & Norbury, 2012; Schraeder, 2008; Ukrainetz, 2015a) | • *A child's performance is compared to normative scores (standard scores means or percentile ranks) derived from a sample of similar peers* |
| **Criterion-referenced:** <br> Assessments that compare a child's performance against a pre-determined level or criterion (i.e. skills expected given a child's age, grade or curriculum level). These assessments may or may not be standardised. <br><br> (Caesar & Kohler, 2009; Paul & Norbury, 2012; Schraeder, 2008; Ukrainetz, 2015a) | • *A child's performance is compared to the curriculum expectations for their year level* <br> • *A child's syntactical and morphological are assessed in relation to knowledge of developmental expectations* |

| | |
|---|---|
| **Descriptive:**<br>Assessments designed to give descriptive or qualitative data on a child's abilities. These assessments may or may not be standardised.<br><br>(Caesar & Kohler, 2009; Paul & Norbury, 2012; Schraeder, 2008; Ukrainetz, 2015a) | • *A child's narrative retell skills are described in terms of strengths and weaknesses*<br>• *A child's social abilities are described in terms of functional abilities observed in the classroom* |
| **Static:**<br>Assessment procedures that are designed to measure performance at a given point in time under given conditions.<br><br>(American Speech-Language-Hearing Association, 2004; Caesar & Kohler, 2009; Dockrell & Marshall, 2015; Kapantzoglou et al., 2012; Leaders Project, 2013; Peña et al., 2006) | • *A child's vocabulary knowledge is assessed in a picture naming task that compares performance to peers of the same age* |
| **Dynamic:**<br>Assessment procedures designed to assess a child's performance under varied conditions or investigate response to intervention. These describe learning potential or identify successful supports and teaching techniques. Includes:<br>- Test-teach-retest procedures<br>- Testing the limits (response to task modification)<br>- Graded levels of prompting (response to different levels of prompting)<br><br>(American Speech-Language-Hearing Association, 2004; Caesar & Kohler, 2009; Dockrell & Marshall, 2015; Kapantzoglou et al., 2012; Leaders Project, 2013; Peña et al., 2006) | • *A child's ability to learn vocabulary is assessed by having the child name a series of pictures, teaching the child the names for pictures they did not know, then retesting using the same pictures to identify response to teaching (test-teach-retest)*<br>• *A teacher re-words or explains questions to determine if poor performance is influenced by not understanding assessment questions; or the teacher modifies the task (such as providing extra visual supports) to compare performance under different conditions (testing the limits)*<br>• *The child's performance on a task is assessed using varied levels of prompting to determine the level or degree of prompting required to learn a skill or successfully complete a task (graded levels of prompting)* |
| **Decontextualised – Hierarchical:**<br>Naturalness of communication:<br>- Discrete or 'pure' skills are assessed, which may be used to infer functional performance.<br>- If conducted by a person: Tasks are directed by the assessor, typically in a 'test' format.<br>- If proxy-reported: Skills, usually skills that the child demonstrates without support, are documented without reference to a specific communicative situation or context.<br>Structure of assessment:<br>- Assessment is highly structured. Each question or item follows on from previous questions or items in a hierarchical (usually developmental) order.<br>- Presentation of subsequent tasks or questions often depends on success with earlier tasks.<br><br>(Koole, Nelson, & Curtis, 2015; Mislevy, Steinberg, & Almond, 2002; Schraeder et al., 1999; Skeat & Perry, 2008; Ukrainetz, 2015b; Westby, 2007) | • *A parent questionnaire asks about the morphological and syntactic abilities that a child demonstrates. Questions are sequenced in order of developmental acquisition, however do not refer to particular communicative situations e.g. Does the child: speak with 3-4 word sentences; use 'ing' verb endings; use 's' regular plural?*<br>• *An SLP assesses morphological and syntactic skills in a series of cloze questions with picture stimulus: "This girl is running, this boy is _____", with questions presented in order of developmental acquisition.*<br>• *A teacher completes a checklist profiling a student's pre-linguistic behaviours at school. Questions are sequenced in developmental order, however do not refer to particular communicative situations e.g. Does the student express pleasure and do they do this through facial expression, body language or gesture?; Does the student request desired items and do they do this through facial expression, body language or gesture?* |

| | |
|---|---|
| **Decontextualised - Non-Hierarchical:**<br>Naturalness of communication:<br>Same as for decontextualised – hierarchical (see above)<br>Structure of assessment:<br> - Questions or tasks are presented one at a time in a structured manner, but do not follow a set hierarchy or sequence (questions could be administered in a different order without consequence).<br> - Questions or items are different from previous questions or items (tasks are not clearly identifiable as following-on from each other).<br><br>(Koole et al., 2015; Mislevy et al., 2002; Schraeder et al., 1999; Skeat & Perry, 2008; Ukrainetz, 2015b; Westby, 2007) | • *A screening checklist asks about behaviours that may indicate language difficulties. Questions are not related to a particular context and are not presented in defined order or sequence e.g. Does the child often: appear to have difficulty thinking of names of objects?; make grammatical errors when speaking?; have difficulty following instructions with 2-3 steps?*<br>• *Knowledge of social rules is assessed through a series of questions that are not related to specific situations in which the child communicates and are not presented in a developmental sequence or hierarchy of difficulty e.g. "What might it mean if someone says "Pull-up your socks?";"What might the doctor say when he greets a patient?"* |
| **Contextualised:**<br>Naturalness of communication<br> - Skills are assessed in a meaningful communicative context. Discrete skills may be targeted, but this occurs within the broader context of a naturalistic communicative situation.<br> - If conducted by a person: Tasks are directed by the assessor but occur in a naturalistic context (e.g. book reading) or a contrived scenario representative of a real situation (e.g. role play). Tasks center on a theme (e.g. a story) with topic continuity across tasks.<br> - If proxy-reported: The child's skills are reported in the context of specific communicative situations or contexts i.e. how does the child communicate in a particular situation.<br>Structure of assessment<br> - Task presentation is less structured and does not typically follow a hierarchical or developmental sequence (as the focus is on meaningful interaction)<br><br>(Koole et al., 2015; Mislevy et al., 2002; Schraeder et al., 1999; Skeat & Perry, 2008; Ukrainetz, 2015b; Westby, 2007) | • *A parent questionnaire assesses communication for different communicative purposes in relation to specific contexts or situations e.g. what does the child do: if they want a toy that is placed out of reach?; when they need to go to the toilet?; if a parent doesn't understand the message they are trying to communicate?*<br>• *Syntactical skills are examined from a transcription of the child recounting their recent trip to the zoo (i.e. microstructure analysis).*<br>• *During interactive book reading activities, an SLP assesses the level of support that a child needs to answer questions. The SLP asks questions about the book using different types of questions and observes the child's response to supports such as repetition of questions and visual prompts*<br>• *A child's ability to respond appropriately to others is observed whilst role-playing real-life scenarios that may occur at school* |
| **Decontextualised – Hierarchical:**<br>Naturalness of communication:<br> - Discrete or 'pure' skills are assessed, which may be used to infer functional performance.<br> - If conducted by a person: Tasks are directed by the assessor, typically in a 'test' format.<br> - If proxy-reported: Skills, usually skills that the child demonstrates without support, are documented without reference to a specific communicative situation or context.<br>Structure of assessment:<br> - Assessment is highly structured. Each question or item follows on from previous questions or items in a hierarchical (usually developmental) order.<br> - Presentation of subsequent tasks or questions often depends on success with earlier tasks.<br><br>(Koole et al., 2015; Mislevy et al., 2002; Schraeder et al., 1999; Skeat & Perry, 2008; Ukrainetz, 2015b; Westby, 2007) | • *A parent questionnaire asks about the morphological and syntactic abilities that a child demonstrates. Questions are sequenced in order of developmental acquisition, however do not refer to particular communicative situations e.g. Does the child: speak with 3-4 word sentences; use 'ing' verb endings; use 's' regular plural?*<br>• *An SLP assesses morphological and syntactic skills in a series of cloze questions with picture stimulus: "This girl is running, this boy is _____", with questions presented in order of developmental acquisition.*<br>• *A teacher completes a checklist profiling a student's pre-linguistic behaviours at school. Questions are sequenced in developmental order, however do not refer to particular communicative situations e.g. Does the student express pleasure and do they do this through facial expression, body language or gesture?; Does the student request desired items and do they do this through facial expression, body language or gesture?* |

**References for Supplementary Appendix 3.1.**

American Speech-Language-Hearing Association. (2004). *Preferred practice patterns for the profession of speech-language pathology [Position Statement]*. Retrieved from [www.asha.org/policy](www.asha.org/policy)

American Speech and Hearing Association. (1993). Language in brief. Retrieved from [http://www.asha.org/Practice-Portal/Clinical-Topics/Spoken-Language-Disorders/Language-In--Brief/](http://www.asha.org/Practice-Portal/Clinical-Topics/Spoken-Language-Disorders/Language-In--Brief/)

Apel, K. (2014). Clinical scientists improving clinical practices: In thoughts and actions. *Language, Speech, and Hearing Services in Schools, 45*, 104-109. doi:10.1044/2014_LSHSS-14-0003

Betz, S. K., Eickhoff, J. K., & Sullivan, S. F. (2013). Factors influencing the selection of standardized tests for the diagnosis of Specific Language Impairment. *Language, Speech and Hearing Services in Schools, 44*, 133-143. doi:10.1044/0161-1461(2012/12-0093)

Beukelman, D. R., & Mirenda, P. (2013). Language development and intervention: Challenges, supports and instructional approaches. In *Augmentative and alternative communication: Supporting children and adults with complex communication needs* (pp. 255-278). Baltimore, MD: Paul H. Brookes Publishing.

Bishop, D. V. M., & McDonald, D. (2009). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of Language & Communication Disorders, 44*(5), 600-615. doi:10.1080/13682820802259662

Caesar, L. G., & Kohler, P. D. (2009). Tools clinicians use: A survey of language assessment procedures used by school-based speech-pathologists. *Communication Disorders Quarterly, 30*(4), 226-236. doi:10.1177/1525740108326334

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. *Child and Adolescent Mental Health, 20*(2), 116-125. doi:10.1111/camh.12072

Eadie, P. (2003). Speech pathology assessment practices: One assessment or many? *Advances in Speech Language Pathology, 5*(1), 65-68. doi:10.1080/14417040510001669081

Edwards, M., Stredler-Brown, A., & Houston, K. T. (2012). Expanding use of telepractice in speech-language pathology and audiology *Volta Review, 112*(3), 227-242.

Hegde, M. N., & Pomaville, F. (2017). Assessment of communication disorders in children. In *Assessment of communication disorders in children: Resources and protocols* (pp. 3-16). Sa Diego, CA: Plural Publishing.

Hyter, Y. D. (2003). Language intervention for children with emotional or behavioral disorders. *Journal of the Council for Children with Behavioral Disorders, 29*(1), 65-76. doi:10.1177/019874290302900104

Kamhi, A. G., Masterson, J. J., & Apel, K. (2007). Assessment of preschool and early school-age children with Developmental Language Disorders. In *Clinical decision-making in Developmental Language Disorders*. USA: Paul. H. Brooks Publishing Co.

Kaminski, R. A., Abbott, M., Aguayo, K. B., Latimer, R., & Good, R. H. (2014). The preschool early literacy indicators: Validity and benchmark goals. *Topics in Early Childhood Special Education, 34*(2), 71-82. doi:10.1177/0271121414527003

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)

Koole, H., Nelson, N. W., & Curtis, A. B. (2015). Factors influencing choices of contextualized versus traditional practices with children and adolescents who have

traumatic brain injury. *Language, Speech, and Hearing Services in Schools, 46*, 352-361. doi:10.1044/2015_LSHSS-14-0109

Larson, V. L., & McKinley, N. L. (2003). Procedures for direct assessment. In *Communication solutions for older students* (pp. 209-268). Eau Claire, WI: Thinking Publications.

Law, J., Campbell, C., Roulstone, S., Adams, C., & Boyle, J. (2008). Mapping practice onto theory: The speech and language practitioner's construction of receptive language impairment. *International Journal of Language and Communication Disorders, 43*, 245-263. doi:10.1080/13682820701489717

Leaders Project. (2013). Understanding assessment: Assessment materials - dynamic vs. static assessment. Retrieved from http://www.leadersproject.org/2013/03/01/assessment-materials-dynamic-vs-static-assessment/

Mashima, P. A., & Doarn, C. R. (2009). Overview of telehealth activities in speech-language pathology. *Telemedicine and e-Health, 14*(10), 1101-1117. doi:10.1089/tmj.2008.0080

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2002). Design and analysis in task-based language assessment. *Language Testing, 19*(4), 477-496. doi:10.1191/0265532202lt240oa

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and Specific Language Impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology, 19*, 78-94. doi:10.1044/1058-0360(2009/09-0028)

Nelson, N. W. (1989). Curriculum-based language assessment and intervention. *Language, Speech, and Hearing Services in Schools, 20*, 170-184. doi:10.1044/0161-1461.2002.170

Newton, P. E. (2007). Clarifying the purposes of educational assessment *Assessment in Education: Principles, Policy & Practice, 14*(2), 149-170. doi:10.1080/09695940701478321

Ockey, G. J. (2009). Developments and challenges in the use of computer-based testing for assessing second language ability. *The Modern Language Journal, 93*(1), 836-847. doi:10.1111/j.1540-4781.2009.00976.x

Olswang, L. B., & Bain, B. A. (1996). Assessment information for predicting upcoming change in language production. *Journal of Speech, Language, and Hearing Research, 39*(2), 414-423. doi:10.1044/jshr.3902.414

Parsons, S., Law, J., & Gascoigne, M. (2005). Teaching receptive vocabulary to children with Specific Ianguage Impairment: A curriculum-based approach. *Child Language Teaching and Therapy, 21*(1), 39-59. doi:10.1191/0265659005ct280oa

Paul, R., & Norbury, C. F. (2012). Chapter 2: Assessment. In R. Paul & C. F. Norbury (Eds.), *Language disorders from infancy through adolescence: Listening, speaking, reading, writing and communicating* (4th ed., pp. 22-60). St Louis, MI: Mosby Elsevier.

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)

Richards, J. A., Xu, D., Gilkerson, J., Yapanel, U., Gray, S., & Paul, T. (2017). Automated assessment of child vocalization development using LENA. *Journal of Speech,*

*Language, and Hearing Research, 60*, 2047–2063. doi:10.1044/2017_JSLHR-L-16-0157

Schraeder, T. (2008). Assessment, evaluation and individualised education programs in schools. In T. Schraeder (Ed.), *A guide to school services in speech-language pathology*. San Diego, CA: Plural Publishing.

Schraeder, T., Quinn, M., Stockman, I. J., & Miller, J. (1999). Authentic assessment as an approach to preschool speech-language screening. *American Journal of Speech-Language Pathology, 8*(3), 195-200. doi:10.1044/1058-0360.0803.195

Serry, T., Rose, M., & Liamputtong, P. (2008). Oral language predictors for the at-risk reader: A review. *International Journal of Speech-Language Pathology, 10*(6), 392-403. doi:0.1080/17549500802056128

Singer, B. D., & Bashir, A. S. (1999). What are executive functions and self-regulation and what do they have to do with language-learning disorders? *Language, Speech, and Hearing Services in Schools, 30*(3), 265-273. doi:10.1044/0161-1461.3003.265

Skeat, J., & Perry, A. (2008). Exploring the implementation and use of outcome measurement in practice: A qualitative study. *International Journal of Language and Communication Disorders, 43*(2), 110-125. doi:10.1080/13682820701449984

Starling, J., Munro, N., Togher, L., & Arciuli, J. (2012). Training secondary school teachers in instructional language modification techniques to support adolescents with language impairment: A randomized controlled trial. *Language, Speech and Hearing Services in Schools, 43*(4), 475-495. doi:10.1044/0161-1461(2012/11-0066)

Ukrainetz, T. A. (2006). Assisting student's in beoming self-regulated writers. In T. A. Ukrainetz (Ed.), *Contextualised language intervention: Scaffolding Pre-K-12 literacy achievement* (pp. 570-572). Eau Claire, WI: Thinking Publications.

Ukrainetz, T. A. (2015a). Contextualised skill intervention framework: The whole and the parts. In T. A. Ukainetz (Ed.), *School-aged language intervention: Evidence-based practices*. US, TX: PRO-ED.

Ukrainetz, T. A. (2015b). The foundations of language intervention: Theory and research. In T. A. Ukainetz (Ed.), *School-age language intervention: Evidence-based practices*. US, TX: PRO-ED.

Vaz, S., Cordier, R., Falkmer, M., Ciccarelli, M., Parsons, R., McAuliffe, T., & Falkmer, T. (2015). Should schools expect poor physical and mental health, social adjustment, and participation outcomes in students with disability? *PLoS ONE, 10*(5), e0126630. doi:10.1371/journal.pone.0126630

Wade, D. T. (2004). Assessment, measurement and data collection tools. *Clinical Rehabilitation, 18*, 233-237. doi:10.1191/0269215504cr183ed

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010a). Assessment of children's literacy via an internet-based telehealth system. *Telemedicine and e-Health, 16*(5), 564-575. doi:10.1089/tmj.2009.0161

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010b). Internet-based telehealth assessment of language using the CELF–4. *Language Speech and Hearing Services in Schools, 41*, 445-458. doi:10.1044/0161-1461(2009/08-0131)

Westby, C. (2007). Application of the ICF in children with language impairments. *Seminars in Speech and Language, 28*(4), 265-272. doi:0.1055/s-2007-986523

Westerveld, M. F., & Claessen, M. (2014). Clinician survey of language sampling practices in Australia. *International Journal of Speech-Language Pathology, 16*(3), 242-249. doi:10.3109/17549507.2013.871336

Williams, C. (2006). Teacher judgements of the language skills of children in the early years

of schooling. *Child Language Teaching and Therapy, 22*(2), 135-154. doi:10.1I

191/0265659006ctQ04oa

Wixson, K. K., & Valencia, S. W. (2011). Assessment in RTI: What teachers and specialists

need to know. *The Reading Teacher, 64*(6), 466-469. doi:10.1598/RT.64.6.13

**Supplementary Appendix 3.2.**

Summary of the Qualitative and Quantitative Data from each Delphi Study Round and the

Changes to the Taxonomy after Round One and Two Data Analysis

| Qualitative data: Themes from comments | Qualitative data: Examples of participant comments related to the identified themes<br><br>R1 = Comment from Round one<br>R2 = Comment from Round two<br>R3 = Comment from Round three<br>NA = Not applicable for this round (as no comments were made related to this theme) | Quantitative data:<br>Level of agreement<br><br>R1: Round one<br>R2: Round two<br>R3: Round three | Changes implemented after each round:<br><br>(Note: no changes after Round three as this was the last round)<br><br>R1 = Changes after Round one<br>R2 = Changes after Round two<br>NA = Not applicable for this round |
|---|---|---|---|
| Aspect I<br>Suggestion to change sequence in flowchart by placing 'comprehension' & production after the other domain categories. | R1: "Consider if the domains should come before *comprehension* and *production*. Much of language requires the integration of *comprehension* and *production* so may be better to consider which domain the child is most challenged in before considering receptive versus expressive (if this is even applicable). Not every language domain has a dominant *comprehension* or *production* component."<br>R2: NA<br>R3: NA | This suggestion was not linked to lack of consensus but was actioned to improve the taxonomy. | R1: Change to the structure of Aspect I so that the components *comprehension* and *production* are placed after other domain categories in the taxonomy flowchart.<br>R2: NA |
| Aspect I<br>Suggestion to add clarification to ensure that categorisation of pre-linguistic communication is clear. | R1: "As the taxonomy is valid for school age children regardless of severity etc, potentially an element that incorporates pre-symbolic and pre-intentional *spoken* language?"<br>R2: NA<br>R3: NA | This suggestion not linked to lack of consensus but was actioned to improve the taxonomy. | R1: Additional information and examples were added to indicate how assessments targeting pre-linguistic communication may be categorised.<br>R2: NA |
| Aspect I<br>Identification of overlap between categories of *discourse* and *social abilities*. | R1: NA<br>R2: "I agree with some definitions for the domains. I do not agree that *discourse* only relates to the types listed, as conversation is a type of discourse, so much of what is classified as *social abilities* is an aspect of discourse."<br>R3: NA | R1 and R2: Many participants selected both (or neither) *discourse* and *social abilities* when describing assessments, indicating potential overlap between these categories. | R1: Additional information was added to the definition of *discourse* and *social abilities* to create greater distinction between these two categories.<br>R2: Amalgamation of discourse and *social abilities* categories into one category. |
| Aspect I<br>Identification of possible overlap between categories *semantics* and *executive functioning* with other categories | R1: NA<br>R2: NA<br>R3: "It is hard to separate the categories of *semantics* and *executive functioning* out as with a case like this as they would likely influence each other". | R1-R3: Lack of consensus on application of components *semantics* and *executive functioning*. | R1: NA<br>R2: NA |
| Aspect I<br>Identification that participants may be considering other possible ways an assessment could be conducted, rather than describing assessments as they were used in case studies. | R1: NA<br>R2: NA<br>R3: "People may choose *semantics* as through language sampling you can calculate TTR [type token ratio] and NDW [number of different words]; however, your case study did not outline this as an analysis used." | R1-R3: Lack of consensus on application of components *semantics* and *executive functioning*. | R1: NA<br>R2: NA |

| Qualitative data: Themes from comments | Qualitative data: Examples of participant comments related to the identified themes<br><br>R1 = Comment from Round one<br>R2 = Comment from Round two<br>R3 = Comment from Round three<br>NA = Not applicable for this round (as no comments were made related to this theme) | Quantitative data: Level of agreement<br><br>R1: Round one<br>R2: Round two<br>R3: Round three | Changes implemented after each round:<br><br>(Note: no changes after Round three as this was the last round)<br><br>R1 = Changes after Round one<br>R2 = Changes after Round two<br>NA = Not applicable for this round |
|---|---|---|---|
| Aspect I Identification that participants may be describing all possible domains, rather than key domains being targeted by the assessment. | R1: "The CELF-4 utilises meta-linguistic skills in the items, though it is not explicitly tested. Working memory is also assessed but I wouldn't classify the CELF4 as assessing broader *executive function*, and the ability to sustain attention is qualitative data obtained from the assessment process."<br>R2: NA<br>R3: "Possibly clinicians thinking more about the secondary skills involved in the questions in the case study e.g. to initiate a conversation with others you need to use *semantic* skills, but there is also an element of forward planning. This I would say is a 'secondary' skill tapped into indirectly - some clinicians might think that the taxonomy factors in these secondary skills." | R1-R3: Lack of consensus on application of the components *semantics* and *executive functioning*. | R1: Additional clarification was added to highlight that domains are only selected if they are specifically targeted and measured by an assessment.<br>R2: Options for this aspect were reduced in the survey to determine if consensus is reached on the main domains (participants could only select one other category in addition to categories that reached consensus in round 2). |
| Aspect II Lack of clarity with *prognostic* categories, particularly the *predict outcome* category. | R1: "I am not sure of any [assessments] in the predict outcome or plan dosage categories"<br>R2: "*Prognostic* tends to lead the reader to the question of whether the young person is likely to improve with or without intervention. *Predict outcome* then tends to make the reader think about this too rather than about supports the young person would need".<br>R3: "*Predict outcome* is not always intuitive to the definition." | R1-R3: Lack of consensus with selection of *prognostic* components to describe assessments | R1: Examples were added to show how *prognostic* categories apply when describing assessments.<br>R2: Examples revised to further highlight application of categories, particularly *predict outcome* category. A name change for the predict *outcome* category was considered, but not implemented due to inability to identify a more suitable name. |
| Aspect II Identification that descriptions of purpose of assessment by be influenced by contextual factors related to service policy (e.g. service policy may assign dosage based on diagnosis rather than response to intervention). | R1 "…the concept of 'dosage' is commonly influenced by many other factors (service restraints, funding, availability)."<br>R2: "I would agree 'specific purpose' section of the assessment purpose, however would rarely separate the *prognostic* and *analytic* areas. Assessment usually requires both areas to be covered at the same time in order to meet the reporting and educational requirements on the service."<br>R3: NA | R1-R3: Lack of consensus with application of aspect II categories to describe assessments. | R1: The assessments being categorised n the Delphi study were placed in into case studies to provide context.<br>R2: Participants were instructed to categorise the assessments in the Delphi study according to the purpose of use in the case study and as though service policy is not an influence. |
| Aspect II Identification that purpose of assessment may be influenced by SLP perspective (e.g. an assessment that is not typically considered *diagnostic* may be used by SLPs in this way; or if SLP views outcome only as change in diagnostic status, then they may identify *detect change* as being the same as *diagnostic*). | R1: "Categorising in this area becomes difficult as the waters easily become muddied between the purpose of the tools (intent/design of the tool) and purpose of use (intent of the examiner). Typical purpose may vary according to clinical context and SLP role"<br>R2: *[Aspect II]* is particularly challenging to categorise, as often this has to do with the nature of the data uncovered and the intent of the clinician in this case.<br>R3: "Perhaps *diagnostic* because some comparison may be made with peers in the mind of the SLP, though the tool as such doesn't make the comparison." | R1-R3: Lack of consensus with selection of aspect II categorises to describe assessments. | R1: The assessments being categorised in the Delphi study were placed in into case studies to provide context.<br>R2: Participants were instructed to categorise the assessments in the Delphi study according to the purpose of use in the case study and as though service policy is not an influence. |
| Aspect II Identification that lack of consensus may arise if participants are considering all possible ways a tool could be | R1: NA<br>R2: NA<br>R3: "Conversation & narrative samples are often analysed using [the Systematic Analysis of Language Transcription] (SALT) database] which does allow for comparison to peers. Some | R1-R3: Lack of consensus with selection of aspect II categorises to describe assessments. | R1: NA<br>R2: NA |

| Qualitative data: Themes from comments | Qualitative data: Examples of participant comments related to the identified themes<br><br>R1 = Comment from Round one<br>R2 = Comment from Round two<br>R3 = Comment from Round three<br>NA = Not applicable for this round (as no comments were made related to this theme) | Quantitative data: Level of agreement<br><br>R1: Round one<br>R2: Round two<br>R3: Round three | Changes implemented after each round:<br><br>(Note: no changes after Round three as this was the last round)<br><br>R1 = Changes after Round one<br>R2 = Changes after Round two<br>NA = Not applicable for this round |
|---|---|---|---|
| used, rather than categorising based only on how assessment is used in the case study. | clinician's may have assumed that [SALT was being used], therefore choosing *diagnostic*". | | |
| Aspect III<br>Lack of clarity with term 'Internet'. | R1: "Examples of internet based are not all using the internet so a possibly confusing term to use if covering other than 'internet'. Would technology or [Information and Communication technologies] *ICT* be better?"<br>R2: NA<br>R3: NA | R1: Lack of consensus with identification of assessments as being able to be conducted via *ICT*. | R1: Change term category name internet to *ICT*.<br>R2: NA |
| Aspect III<br>Lack of clarity with structure of aspect III in the taxonomy. | R1: "...if you have two areas - delivery and setting why you don't have a box with these labelled in between the Aspect III box & the 8 boxes divided into the 2 categories?"<br>R2: NA<br>R3: NA | R1-R3: Lack of consensus across Aspect III. | R1: Change to the structure of Aspect III to show a component for *method* and a component for *environmental context*.<br>R2: NA |
| Aspect III<br>Identification that lack of consensus may arise from differences between purposes for which assessments are used due to differences in SLP perspective. | R1: "These responses reflect my use of the CELF-4 only and do not necessarily encompass how else the test may be delivered."<br>R2: NA<br>R3: NA | R1-R3: Lack of consensus across Aspect III. | R1: The assessments being categorised in the Delphi study were placed in into case studies to provide context.<br>R2: NA |
| Aspect III<br>Lack of clarity with definition of *software*. | R1: "Computer programs and Apps play an important role in language sample analysis, but do not deliver the assessment, as such. Similarly, the CCC-2 can be scored using software, but is not delivered in this way."<br>R2: NA<br>R3: NA | This suggestion not linked to lack of consensus. | R1: Additional clarification added to explain that the term *software* only applies when the assessment is primarily delivered by a software program.<br>R2: NA |
| Aspect III<br>Lack of clarity with definitions for *environmental context* with some participants interpreting this as being physical location. | R1: "Assessment may be conducted in the clinic or school but draw on child performance in another setting such as home or community. The definitions may then be unclear/confusing"<br>R2: "Difficulty in relation to [case study two] and describing environment. Seen at school but in a withdrawal situation which more closely resembles clinic than classroom environment"<br>*R3: "*Perhaps it's due to an intuitive level of response - as the interview was conducted in the clinic although [it] is a proxy report. Maybe it's just hard to tick home when the interview is in the clinic?" | R1-R3: Lack of consensus with selection of *environmental context* categories to describe assessments. | R1: Changed category name from setting to *environmental context* to highlight that *environmental context* is being described (not physical location). Clarification and examples added to category definition to highlight that *environmental context* is being described (not physical location).<br>R2: Further clarification added to highlight that the category describes *environment context* (not physical location). |
| Aspect III<br>Lack of clarity with definitions for *school context* with some participants focussing on one element in the assessment, rather than categorising based on the category that best describes the assessment overall. | R1: NA<br>R2: NA<br>R3: "While the assessment is conducted at school it is in a withdrawal/clinical setting. The fact that part of the protocol is that the student brings a piece of school work to share and discuss in the conversational element may lead to confusion." | R1-R3: Lack of consensus with identification of assessments in environmental context. | R1: NA<br>R2: NA |
| Aspect III<br>Lack of clarity with *environmental context* with some participants | R1: NA<br>R2: "...*clinical assessment* might be better described as *de-contextualised* (i.e. .focus is on the within-person skills assessed separate from partners | R1-R3: Lack of consensus with identification of assessments in | R1: NA<br>R2: Additional information added to highlight that Aspect III *environmental context* |

| Qualitative data: Themes from comments | Qualitative data: Examples of participant comments related to the identified themes<br><br>R1 = Comment from Round one<br>R2 = Comment from Round two<br>R3 = Comment from Round three<br>NA = Not applicable for this round (as no comments were made related to this theme) | Quantitative data: Level of agreement<br><br>R1: Round one<br>R2: Round two<br>R3: Round three | Changes implemented after each round:<br><br>(Note: no changes after Round three as this was the last round)<br><br>R1 = Changes after Round one<br>R2 = Changes after Round two<br>NA = Not applicable for this round |
|---|---|---|---|
| confusing the aspect III distinction with *environmental context* with the Aspect IV distinction *task type*. | and environment where communication occurs) and community might be better described as *contextualised* (i.e. focus is on the within-person skills assessed within naturalistic interactions with partners in the environment where communication occurs)" | *environmental context...* | identifies the environment in which skills are being assessed and Aspect IV *task type* identifies the communicative tasks used in the assessment. |
| Aspect III<br>Lack of clarity between *proxy-reported* vs. *conducted by SLP* with some participants confusing SLP actions (e.g. interviewing a parent) with method by which data is collected (e.g. parent reports information). | R1: "I found the terms indirect and reported were confusing."<br>R2: "Could a *proxy report* still be recorded in the moment? e.g.: behavioural observation writing down exactly what occurs & this is then reviewed at a later date?" [Note: The behavioural observation described by this participant would be considered assessment conducted by a person and not information obtained through *proxy-report*].<br>R3: NA | R1-R2: Lack of consensus with identification of parent interview/ questionnaires as *proxy-reported* or *conducted by SLP*. | R1: Removal of terms 'direct' and 'indirect' from category names. Restructuring of categories within Aspect III to better represent distinctions between categories<br>R2: Additional information added to highlight the difference between *proxy-reported* and *conducted by SLP*. |
| Aspect III<br>Identification that applying the taxonomy to describing case studies may require a high level of information processing. | R1: NA<br>R2: NA<br>R3: "Participants have not read the definitions (and associated examples) properly." | R1-R3: Lack of consensus with identification of assessments across Aspect III. | R1: NA<br>R2: NA |
| Aspect IV<br>Identification that lack of consensus may arise if participants are considering all possible ways an assessment could be used, rather than categorising based only on how assessment is used in the case study. | R1: "The language sampling protocol can be *norm-referenced* but only if there is a relevant/appropriate database."<br>R2: "I found '*descriptive*' tricky [to identify] with reference to the narrative assessment. They are and can be criterion referenced as well."<br>R3: SLPs might not be familiar enough with the language sampling protocol to know that it is somewhat *standardised* - often narrative & conversation samples are thought of (& conducted) in a less structured way. | R1-R3: Lack of consensus with selection of Aspect IV categories to describe assessments. | R1: The assessments being categorised in the Delphi study were placed in into case studies to provide context. Additional information added to highlight that categories are selected based on how assessments are used in case studies.<br>R2: Further explanation that categories are selected based on how assessments are used in case studies. |
| Aspect IV<br>Identification that *task-type* categories *contextualised* and *activity-focused* may be difficult to distinguish between. | R1: "In theory, the definitions were clear, however I found the checklists more challenging to rate based on the definitions between *contextualised* and *activity focused*"<br>R2: "Decision making regarding *contextualised* and *activity-focused* [is] not always clear."<br>R3: "*Contextualised* and *activity-focused* categories overlap to an extent." | R1-R3: Lack of consensus with selection of Aspect IV *task type* categories to describe assessments. | R1: Revision of definitions and examples added to highlight key distinctions between *task type* categories.<br>R2: Information on the definitions was formatted under headings to assist with application of terms. |
| Aspect IV<br>Identification that lack of consensus may arise if SLPs apply definitions that are different to definitions in the taxonomy. | R1: NA<br>R2: NA<br>R3: "Possibly [confusion] in regards to my understanding of *dynamic* assessment? It seems clear in your definition however." | R1-R3: Lack of consensus with identification assessments in case study two as *standardised* and *dynamic* | R1: NA<br>R2: NA |
| Aspect IV<br>Identification that applying the taxonomy to describing case studies may require a high level of information processing. | R1:NA<br>R2:NA<br>R3: "The definitions contain a lot of detail which is hard to hold on to when flipping back [through the reference document] to think about what was done in the assessment." | R1-R3: Lack of consensus with selection of Aspect IV *task type* categories to describe assessments. | R1: NA<br>R2: NA |
| Aspect I-IV<br>Identification that applying the taxonomy | R1: NA<br>R2: NA | R1-R3: Lack of consensus on application of some | R1: NA<br>R2: NA |

| Qualitative data: Themes from comments | Qualitative data: Examples of participant comments related to the identified themes<br><br>R1 = Comment from Round one<br>R2 = Comment from Round two<br>R3 = Comment from Round three<br>NA = Not applicable for this round (as no comments were made related to this theme) | Quantitative data: Level of agreement<br><br>R1: Round one<br>R2: Round two<br>R3: Round three | Changes implemented after each round:<br><br>(Note: no changes after Round three as this was the last round)<br><br>R1 = Changes after Round one<br>R2 = Changes after Round two<br>NA = Not applicable for this round |
|---|---|---|---|
| to describing case studies may require a high level of information processing as this may have influenced application. | R3: "The amount of information needed to be taken into account in the case studies [may influence application]." | components of the taxonomy. | |
| Overall Taxonomy Participants identified as finding the taxonomy useful for conceptualising clinical work. | R1: "I really like this classification. I use most if not all types of assessment but had never considered the different types so explicitly. I think it will add hugely to professional education at [universities] and work places to help build a more conscious and explicit awareness of what we do."<br>R3: "I think it's a great classification and useful." | NA | NA |
| Overall Taxonomy Participants identified that understanding and applying the taxonomy accurately takes time and consideration. | R2: "Challenging to keep all parameters in mind. I hope I have not been too hasty in my responses."<br>R3: "I think the assessment type classification is complex and a new way of thinking. [It] takes real consideration to use." | NA | NA |
| Overall taxonomy Participants commented that the taxonomy and their understanding of the taxonomy improved over rounds and that examples assisted in improving the taxonomy. | R2: "The definitions were helpful in considering the options."<br>R3: "The new additions to definitions and examples have helped clarify the taxonomy." | NA | NA |

**Supplementary Appendix 3.3.**

Delphi Study Round One Survey Questions


**I consent to answering questions in an on-line survey and for my responses to be used for the purposes described above**

*[Yes/No response. If no, skip to end of survey]*


**ELLIGIBILITY TO PARTICIPATE**

**Do you have (or are eligible for) certified practicing membership with Speech Pathology Australia?**

*[Yes/No response. If no, skip to end of survey]*

**Have you spent more than 5 years (full-time equivalent) in the last 10 years engaged in assessment, intervention, education or research activities related to students aged 4-18 years with language disorder?**

For the purpose of this study:

"Students with language disorder" refers to children and adolescents with oral or written language support needs (i.e. semantics, syntax, morphology, phonology, discourse or pragmatics) regardless of primary diagnosis, severity, aetiology or other co-morbidities associated with the language support needs. The focus of this study is mono-lingual English-speaking students.

"Activities" include:

a) Provision of clinical services (where approximately 50% or more of caseload is students aged 4-18 years with language disorder).

b) Research (where approximately 50% or more of research activities relate to students aged 4-18 years with language disorder).

c) Professional supervision/support, academic teaching, resource development or consultancy (where approximately 50% or more of professional activities relate to services for children aged 4-18 years with language disorder).

d) Combination of the above.

*[Yes/No response. If no, skip to end of survey]*


**PARTICIPANT DEMOGRAPHICS**

Please indicate the option(s) that best describe the sector(s) in which you are currently employed as a speech pathologist (or in other work related to child language development or education). Select a maximum of 2 options.

*[Multiple choice response]*

**Please indicate your (completed) qualifications. Note: It is not necessary to indicate qualifications that are unrelated to speech pathology, child development or education.**

*[Multiple choice response]*

**Please indicate the number of years in total (full-time equivalent) that you have worked as a speech pathologist (or in other employment related to child language development or education).**

*[Multiple choice response]*

For the remaining questions on this survey, you will need to refer to the document in the following link:

[Delphi Study Reference Sheet](#)

**ASPECT I**

Please refer to the document titled: <u>Delphi Study Reference Sheet</u>. Consider the information presented regarding the structure of Aspect I (Language Domain). The categories in this aspect are not mutually exclusive (i.e. assessments and interventions may target multiple domains).

**Overall, the structure of Aspect I seems useful for describing the broad target areas for spoken language assessments and interventions for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate what changes you would make to the structure of Aspect I (Language Domain) and where possible, provide references or reasoning.**

*[Open answer response]*

**Do you agree with the definitions provided for the components of Aspect I (Language Domain)?**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate what changes you would make to the definitions for Aspect I (Language Domain) and where possible, provide references.**

*[Open answer response]*

To examine the usefulness of the proposed taxonomy for classifying assessments in a meaningful and consistent way, you are now asked to consider the following assessments and how they would be categorised according to Aspect I (Assessment Domain) in its current

form. If you wish to see further information, click on the assessment names for links to information from websites about each assessment (note: it is not a requirement that you read all the information in these links).

Clinical Evaluation of Language Fundamentals CELF-4 (Semel et al., 2006)

Children's Communication Checklist 2nd Edition CCC-2 (Bishop, 2003)

Language Sampling Protocol (Westerveld & Gillon, 2002)

Pragmatics Profile for Children (Dewart & Summers, 1995)

If you do not feel that you know a particular assessment well enough to categorise it, then click in column one ("unfamiliar") for that particular assessment and do not complete other columns.

If you are familiar with the assessment then leave column one blank and select answers from the other columns. Refer to the information in the Delphi Study Reference Sheet when categorising. If unsure about any answers, then try to select the option/s that you think best fit.

**Please categorise the following assessments according to Aspect I (Assessment Domain) of the proposed taxonomy.**

*[Closed choice answer]*

**If you have any comments about Aspect I (Assessment Domain) or the categorisation of assessments within this aspect, please comment here.**

*[Open answer response]*

**ASPECT II**

Please refer to the document titled Delphi Study Reference Sheet. Consider the information regarding the structure and definitions for Aspect II (Assessment Purpose).

Assessments may have more than one purpose, however when categorising, the intended (or typical) purposes of an assessment are identified, rather than all the possible purposes that an assessment could or might have.

**Overall, the structure of Aspect II (Assessment Purpose) seems useful for describing the purposes of language assessments for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate what changes you would make to the structure of Aspect II (Assessment Purpose) and, where possible, provide references or reasoning.**

*[Open answer response]*

**Do you agree with the definitions provided for Aspect II (Assessment Purpose)?**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate what changes you would make to the definitions for Aspect II (Assessment Purpose) and, where possible, provide references.**

*[Open answer response]*

You are now asked to categorise assessments according to Aspect I (Assessment Domain) in its current form (links to information on the assessments and the Delphi Study Reference sheet are provided again).

**Please categorise these assessments according to Aspect II (Assessment Purpose) of the proposed taxonomy.**

*[Closed choice answer]*

**If you have any comments about Aspect II (Assessment Purpose) or the categorisation of assessments within this aspect, please comment here.**

*[Open answer response]*

**ASPECT III**

Please refer to the document titled <u>Delphi Study Reference Sheet</u>. Consider the information regarding the structure and definitions for Aspect III (Assessment Purpose). It is possible that an assessment is able to be conducted by both an SLP and another person and conducted both face-face and using internet.

Assessments may be conducted in more than one setting, however when categorising, consider the settings in which the assessment is most typically or appropriately used, rather than all the settings in which it could or might be used.

**Overall, the structure of Aspect III (Assessment Delivery) seems useful for describing the delivery of language assessments for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate what changes you would make to the structure of Aspect III (Assessment Delivery) and, where possible, provide references or reasoning.**

*[Open answer response]*

**Do you agree with the definitions provided for Aspect III (Assessment Delivery)?**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate what changes you would make to the definitions for Aspect III (Assessment Delivery) and, where possible, provide references.**

*[Open answer response]*

You are now asked to categorise the assessments according to Aspect III in its current form (links to information on the assessments and the Delphi Study Reference sheet were provided again).

**Please categorise these assessments according to Aspect III (Assessment Delivery) of the proposed taxonomy.**

*[Closed choice answer]*

**If you have any other comments about Aspect III (Assessment Delivery) or the categorisation of assessments within this aspect, then please comment here.**

*[Open answer response]*

**ASPECT IV**

Please refer to the document titled: Delphi Study Reference Sheet. Consider the information regarding the structure and definitions for Aspect IV A (Assessment Form).

**Overall, the structure of Aspect IV (Assessment Form) seems useful for describing the different forms of language assessments for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate what changes or additions you would make to the structure of Aspect IV (Assessment Form) and, where possible, provide references or reasoning.**

*[Open answer response]*

**Do you agree with the definitions provided for Aspect IV (Assessment Form)?**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree",*

*"Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree",*

*"Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate what changes you would make to the definitions for Aspect IV**

**(Assessment Form) and, where possible, provide references.**

*[Open answer response]*

You are now asked to categorise the assessments according to Aspect IV in its current form

(links to information on the assessments and the Delphi Study Reference sheet are provided

again).

**Please categorise these assessments according to the standardisation, data type and the**

**static/dynamic distinctions described in Aspect IV (Assessment Form), of the proposed**

**taxonomy.**

*[Closed choice answer]*

**If you have any other comments about any of the components in Aspect IV (Assessment**

**Form) or the categorisation of assessments within this aspect, then please comment**

**here.**

*[Open answer response]*

**OVERALL TAXONOMY STRUCTURE**

You are now asked your opinion on the overall structure of the taxonomy (i.e. number of

aspects and sequence or layout of aspects). Refer to the document titled Delphi Study

Reference Sheet.

**The overall structure of the taxonomy seems useful for describing assessments and interventions for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please comment on what you would add, remove or change with regards to the overall structure of the taxonomy. Where possible, provide references or reasoning.**

*[Open answer response]*


**FINAL COMMENTS**

**Do you have any other comments or feedback regarding this proposed taxonomy that have not been provided elsewhere? If so, please write here.**

*[Open answer response]*

**Supplementary Appendix 3.4.**

Delphi Study Round Two Survey Questions

**I consent to answering questions in an on-line survey and for my responses to be used for the purposes described above**

*[Yes/No response. If no, skip to end of survey]*

**ELIGIBILITY TO PARTICIPATE**

Only participants who completed round one (i.e. progressed to the last page with the statement "Thank-you for completing this survey") are able to complete round two. This is because the content of round two requires participants to have the background information from round one.

**Did you complete the Round One survey in this Delphi Study?**

*[Yes/No response]*

**PARTICIPANT DEMOGRAPHICS**

[Please see questions from Delphi Study Survey Round One]

The following question asks you to provide your email address. This question is optional.

**Please provide your email address here:**

*[Open answer response]*

Please open the document in this link: Delphi Study Feedback Sheet R2

This document summarises the results of round one and explains the content of round two.

Whilst you do not have to read all the details in the tables, it is important that you understand the findings from round one and the aims of round two.

Now, please open the document in this link: Delphi Study Reference Sheet v2.

You will need to refer to this document whilst completing the questions in this survey. This document is the same as the document for Round One, with changes/additions indicated in red font. You do not have to read this entire document; however, you do need to read and consider the changes indicated in red font.

**ASPECT I**

Please refer to the document in the link: Delphi Study Reference Sheet v2. Consider the information regarding the structure and definitions for Aspect I (Assessment Domains).

**Overall, the structure of Aspect I seems useful for describing the broad target areas for language assessments and interventions for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate specifically what you would change regarding the structure of Aspect I (Domain). Provide references or reasoning for your suggested changes.**

*[Open answer response]*

**Do you agree with the definitions provided for the components of Aspect I (Language Domain)?**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Where you do not agree, please provide alternative definition/s for categories in Aspect I (Domain)**

*[Open answer response]*

**ASPECT II**

Please refer to the document titled <u>Delphi Study Reference Sheet v2.</u> Consider the information regarding the structure and definitions for Aspect II (Assessment Purpose).

Assessments may have more than one purpose, however when categorising, the intended (or typical) purposes of an assessment are identified, rather than all the possible purposes that an assessment could or might have.

**Overall, the structure of Aspect II (Assessment Purpose) seems useful for describing the purposes of language assessments for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate specifically what you would change regarding the structure of Aspect II (Assessment Purpose). Provide references or reasoning for your suggested changes.**

*[Open answer response]*

**Do you agree with the definitions provided for Aspect II (Assessment Purpose)?**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Where you do not agree, please provide alternative definition/s for categories in Aspect II (Assessment Purpose)**

*[Open answer response]*

**ASPECT III**

Please refer to the document titled Delphi Study Reference Sheet v2. Consider the information regarding the structure and definitions for Aspect III (Assessment method and delivery).

Assessments are either conducted in person, via software or via proxy-report. Assessments conducted by a person may be conducted by an SLP or another person. They may also be conducted face-face or via ICT.

When categorising based on environment, consider the environmental context being assessed (which may not be the same as a physical location).

**Overall, the structure of Aspect III (Assessment Delivery) seems useful for describing the delivery of language assessments for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate specifically what changes you would make to the structure of Aspect III (Assessment Delivery). Provide references or reasoning for your suggested changes.**

*[Open answer response]*

**Do you agree with the definitions provided for Aspect III (Assessment Delivery)?**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Where you do not agree, please provide alternative definition/s for categories in Aspect III (Assessment Delivery).**

*[Open answer response]*

**ASPECT IV**

Please refer to the document titled Delphi Study Reference Sheet v2. Consider the information regarding the structure and definitions for Aspect IV (Assessment Form).

**Overall, the structure of Aspect IV (Assessment Form) seems useful for describing the different forms of language assessments for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate specifically what changes or additions you would make to the structure of Aspect IV (Assessment Form). Provide references or reasoning for your suggested changes.**

*[Open answer response]*

**Do you agree with the definitions provided for Aspect IV (Assessment Form)?**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Where you do not agree, please provide alternative definition/s for categories in Aspect IV (Assessment Form)**

*[Open answer response]*

**OVERALL TAXONOMY STRUCTURE**

You are now asked your opinion on the overall structure of the taxonomy (i.e. number of aspects and sequence or layout of aspects).

Refer to the document titled Delphi Study Reference Sheet v2.

**Overall, the structure of the taxonomy seems useful for describing assessments and interventions for school aged children.**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please comment specifically on what you would add, remove or change with regards to the overall structure of the taxonomy. Provide references or reasoning for your suggested changes.**

*[Open answer response]*

**FINAL TAXONOMY COMMENTS**

**If you have any other comments (not provided previously) about any aspect of the taxonomy, please write here.**

*[Open answer response]*

**ASSESSMENT CASE STUDIES**

You are now asked to consider two short case studies, each describing an assessment for a school-aged student with language disorder. You will be asked to describe the assessment in each case study according to the proposed taxonomy.

You do not need to be familiar with the assessments to complete the questions, as all the information is provided in the case study. In fact, we ask that you do not consider information that is not given in the case study. The purpose is to determine if language experts apply the taxonomy in the same way when categorising from the same information. Even if you think of different ways that the assessments could be conducted; or even if you conduct these assessments differently yourself, please only categorise based on how the assessment is conducted in the case study.

Note: These case studies were created for the purpose of this Delphi Study. They have been kept succinct (for the ease of Delphi Study participants) and are not intended to be fully comprehensive descriptions of an assessment process. They are not intended to be examples of "recommended practice" nor are they intended to represent how assessments are most frequently used in SLP practice.

**Please describe the assessments in the following case studies according to Aspect I (Assessment Domain) of the proposed taxonomy**. When answering, refer to the Delphi Study Reference Sheet v2.

Click on the links below to open the assessment case studies:

Assessment Plan for Meg (Pragmatics Profile)

Assessment Plan for Eric (Language Sampling Protocol)

*[Closed choice answer with Aspect I taxonomy categories]*

**Please describe the following assessments according to Aspect II (Assessment Purpose) of the proposed taxonomy** (links to information on the assessments and the Delphi Study Reference sheet were provided again).

*[Closed choice answer with Aspect II taxonomy categories]*

**Please describe the following assessments according to Aspect III (Assessment Delivery) of the proposed taxonomy** (links to information on the assessments and the Delphi Study

Reference sheet were provided again).

*[Closed choice answer with Aspect III taxonomy categories]*

**Please describe the following assessments according to Aspect IV (Assessment Form) of the proposed taxonomy** (links to information on the assessments and the Delphi Study Reference sheet were provided again).

*[Closed choice answer with Aspect III taxonomy categories]*

**FINAL COMMENTS ON CASE STUDIES**

If you have any comments about the taxonomy for describing assessments (either the assessments in the case studies or other assessments), then please comment here.

*[Open answer response]*

**Supplementary Appendix 3.5.**

Delphi Study Round Three Survey Questions

**I consent to answering questions in an on-line survey and for my responses to be used for the purposes described above**

*[Yes/No response. If no, skip to end of survey]*

**ELIGIBILITY TO PARTICIPATE**

Only participants who completed round one (i.e. progressed to the last page with the statement "Thank-you for completing this survey") are able to complete round two. This is because the content of round two requires participants to have the background information from round one.

**Did you complete the Round One survey in this Delphi Study?**

*[Yes/No response]*

**PARTICIPANT DEMOGRAPHICS**

[Please see questions from Delphi Study Survey Round One]

The following question asks you to provide your email address. This question is optional.

**Please provide your email address here:**

*[Open answer response]*

**ASPECT IA AMMENDMENT**

The aspect I categories "Social Abilities" and "Discourse" were merged into a single category called "Social-Abilities & Discourse". This change was made to address difficulties in defining two distinctive, mutually exclusive categories (i.e. to address overlap between the

two categories). The definitions within these categories are largely unchanged; however, as this is structural change to the taxonomy, participants are asked to indicate their level of agreement with the merger (List of changes was provided in the survey).

**Please indicate your level of agreement with the merged category "Social-Abilities & Discourse".**

*[Likert scale response: "Strongly Agree", "Agree", "Neither agree or disagree", "Disagree", "Strongly Disagree"]. If participant answers "Neither agree or disagree", "Disagree" or "Strongly Disagree" then direct to following question:*

**Please indicate why you do not agree with the category "Social Abilities & Discourse":**

*[Two option response: 1. "I prefer the two separate categories of "Social Abilities" and "Discourse" (i.e. as they were in round two)" or 2. Other reason. Please specify: _____ ]*

**ASSESSMENT CASE STUDIES**

The next part of the survey asks you categorise the same assessment case studies from round two (with only very minor adjustments) on the categories that were not agreed upon in round two. You do not need to be familiar with the assessments in the case studies in order to describe them using the taxonomy. The purpose is to determine if language experts apply the taxonomy in the same way when categorising from the same information. Therefore, even if you think of different ways that the assessments could be conducted; or even if you conduct these assessments differently yourself, it is important that you only categorise based on how the assessment is conducted in the case study.

Note: These case studies were created for this Delphi Study. They are not intended to be examples of "recommended practice" nor are they intended to represent how assessments are most frequently used in SLP practice.

Links for case studies:

Assessment Plan for Meg (Pragmatics Profile)

Assessment Plan for Eric (Language Sampling)

Read the case studies and the category definitions provided in the tables below, then answer the questions.

If you wish to see the reference list, or read the background information for any of the definitions, then please refer to the Delphi Study Reference Sheet v3

Aspect I Case 1 (Meg). Assessment Domain.

In round two, participants:

Agreed that "Spoken Language", "Comprehension", "Production" and "Social Abilities/Discourse" apply to this case study

Agreed that "Written Language", "Morphosyntax" and "Meta-Abilities" do not apply to this case study

Participants did not agree on categories "Semantics" and "Executive Functions" for case study 1 (definitions for these categories provided in the survey as well as the links to case studies and the Delphi Study Reference Sheet v3).

**Aspect I case 1 (Meg). Please indicate if you think one of these categories describes case study 1:**

*[Closed choice answer: "Semantics", "Executive Functions" or "Neither of these"]*

**Aspect I Case 1 (Meg). If the components "Semantics" and "Executive Functioning" do not reach consensus for case study 1 during round three, what do you think would be the reason? (select one answer)**

*[Closed choice options:*

*There is overlap between categories in this aspect, which makes categorisation difficult. If so,*

*please indicate which categories overlap:  _____*

*Category definition/s lack clarity or may be open to misinterpretation. If so, please indicate*

*which definitions are unclear:  _____*

*Category name/s are used differently in other literature, which may cause misinterpretation*

*when applying taxonomy. If so, please indicate which category name/s are open to*

*misinterpretation:  _____*

*The case study lacks information needed to categorise. If so, please indicate what*

*information is lacking:  _____*

*Don't know why there would be lack of consensus for these components*

*Other reason. Please specify:  _____ ]*

Aspect I Case 2 (Eric). Assessment Domain.

In round two, participants:

Agreed that "Spoken Language", "Comprehension", "Production", "Morphosyntax" and

"Social Abilities/Discourse" apply to this case study

Agreed that "Written Language" and "Meta-Abilities" do not apply to this case study

Participants were not in agreement with regards to categories "Semantics" and "Executive

Functioning" (definitions for these categories provided in the survey as well as the links to

case studies and the Delphi Study Reference Sheet v3).

**Aspect I Case 2 (Eric).  Please indicate if you think one of these categories describes case**

**study 2:**

*[Closed choice answer: "Semantics", "Executive Functions" or "Neither of these"]*

**If the components "Semantics" and "Executive Functioning" do not reach consensus**

**for case study 2 (Eric) during round three, what do you think would be the reason?**

**(select one answer)**

Please see closed choice options for Aspect I Case 1 (Meg)

Aspect II Case 1 (Meg). Prognostic Purposes.

In round two, participants:

Agreed that "Select Intervention" applies to this case study.

Participants were not in agreement with regards to categories "Predict Outcome" and "Plan Dosage" (definitions for these categories provided in the survey as well as the links to case studies and the Delphi Study Reference Sheet v3).

**Aspect II (Prognostic) Case 1 (Meg). Please indicate if you think one of these categories describes case study 1:**

*[Closed choice answer: "Predict Outcome", "Plan Dosage" or "Neither of these"]*

Aspect II Case 1 (Meg). Analytical Purposes

In round two, participants:

Agreed that "Describe Status" applies to this assessment and agreed that "Detect Change" does not apply to this case study.

Participants were not in agreement with regards to categories "Screening" and "Diagnosis" (definitions for these categories provided in the survey as well as the links to case studies and the Delphi Study Reference Sheet v3).

**Aspect II (Analytical) Case 1 (Meg). Please indicate if you think one of these categories describes case study 1.**

*[Closed choice answer: "Screening", "Diagnosis" or "Neither of these"]*

**If the purposes "Predict Outcome", "Plan Dosage", Screening" and "Diagnosis" do not reach consensus for case study 1 during round three, what do you think would be the**

**reason? (select one answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

Aspect II Case 2 (Eric). Prognostic Purposes.

Participants did not agree on any "prognostic" categories for this case study.

Participants were not in agreement with regards to categories "Predict Outcome", "Select

Intervention", "Plan Dosage" (definitions for these categories provided in the survey as well

as the links to case studies and the Delphi Study Reference Sheet v3).

**Aspect II Prognostic Case 2 (Eric). Please indicate if you think one of these categories**

**describes case study 2:**

*[Closed choice answer: "Predict Outcome", "Select Intervention", "Plan Dosage", "None*

*of these]*

Aspect II Case 2 (Eric). Analytical Purposes.

In round two, participants:

Agreed that "Describe Status" applies to this assessment.

Agreed that "Screening" does not apply to this case study.

Participants were not in agreement with regards to categories "Detect Change" and

"Diagnostic" (definitions for these categories provided in the survey as well as the links to

case studies and the Delphi Study Reference Sheet v3).

**Aspect II Analytical Case 2 (Eric). Please indicate if you think one of these categories**

**describes case study 2:**

*[Closed choice answer: "Diagnostic", "Detect Change", "Neither of these]*

**If the purposes "Predict Outcome", "Select Intervention", "Plan Dosage",**

**"Diagnostic" and "Detect Change" do not reach consensus for case study 2 during**

**round three, what do you think would be the reason? (select one answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

Aspect III Case 1 (Meg). Assessment Delivery.

In round two, participants:

Disagreed regarding method i.e. whether case study is "Conducted by SLP" or Proxy-Report"

(Definitions for these categories provided in the survey as well as the Delphi Study Reference

Sheet v3).

Disagreed regarding environment i.e. whether case study is "Clinical" or "Community-

Home" (definitions for these categories provided in the survey as well as the links to case

studies and the Delphi Study Reference Sheet v3).

**Aspect III Case 1 (Meg). Please select the categories that you think describe case study**

**1:**

**Method** *[Closed choice answer: "Conducted by SLP", "Proxy Report"]*

**Environment** *[Closed choice answer: "Clinic", "Community-Home"]*

**If the components "Conducted by SLP" and "Proxy-report" do not reach consensus for**

**case study 1 (Meg) during round three, what do think would be the reason? (select one**

**answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

**If the components "Clinic" and "Home" do not reach consensus for case study 1 (Meg)**

**during round three, what do think would be the reason? (select one answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

Aspect II Case 2 (Eric). Assessment Delivery.

In round two, participants:

Agreed that assessment is "Conducted by SLP."

Disagreed regarding environment i.e. whether case study is "Clinical" or "Community-School" (definitions for these categories provided in the survey as well as the links to case studies and the Delphi Study Reference Sheet v3).

**Aspect III Case 2 (Eric). Please the category that you think describes case study 2: Environment** *[Closed choice answer: "Clinic", "Community-School"]*

**If the categories "Clinic" and "School" do not reach consensus for case study 2 (Eric) in round three, what do you think will be the reason? (select one answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

Aspect IV Case 1 (Meg). Assessment Form.

In round two, participants:

Agreed that categories "Non-Standardised", "Static" and "Descriptive Data" describe this assessment

Disagreed with regards to task-type (definitions for these categories provided in the survey as well as the links to case studies and the Delphi Study Reference Sheet v3).

**Aspect IV Case 1 (Meg). Please select the category that you think describes case study 2: Task-Type** *[Closed choice answer: "Hierarchical", "Non- hierarchical" "Contextualised" "Activity Focused"]*

**If task-type does not reach consensus for case study 1 (Meg) during round three, what do you think would be the reason? (select one answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

Aspect IV Case 2 (Eric). Assessment Form.

In round two, participants:

Agreed on "Descriptive data"

Disagreed with regards to "Standardised" vs "Non-Standardised"; "Static" vs "Dynamic" and

"Task Type" (definitions for these categories provided in the survey as well as the links to

case studies and the Delphi Study Reference Sheet v3).

**Aspect IV Case 2 (Eric). Please select the categories that you think describe case study**

**2:**

**Standardisation** *[Closed choice answer: "Standardised", "Non-Standardised"]*

**Static or dynamic** *[Closed choice answer: "Static", "Dynamic"]*

**Task-Type** *[Closed choice answer: "Hierarchical", "Non- hierarchical" "Contextualised"*

*"Activity Focused"]*

**If the components "Non-standardised" and "Standardised" do not reach consensus for**

**case study 2 (Eric) during round three, what do you think would be the reason? (select**

**one answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

**If the components "Static" and "Dynamic" do not reach consensus for case study 2**

**(Eric) during round three, what do you think would be the reason? (select one answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

**If task-type does not reach consensus for case study 2 (Eric) during round three, what**

**do you think would be the reason? (select one answer)**

[Please see closed choice options for Aspect I Case 1 (Meg)]

**FINAL COMMENTS**

**If you have any other comments or feedback regarding the taxonomy for describing assessments (either the case studies or other assessments), then please comment here.**

*[Open answer response]*

**Supplementary Appendix 3.6.**

Case Study One (Assessment Plan)

| Background Information | |
|---|---|
| **Student's name** | Meg |
| **Student's Age** | 4;10 years |
| **Summary of existing information** | Currently about to start first year of formal schooling (i.e. Kindy/Prep). Has a history of delayed/atypical communication development, repetitive behaviours and sensory processing problems. Diagnosed with Autism Spectrum Disorder at 3;0 years (previous reports available from paediatrician and therapists). |
| | Currently speech is limited to occasional delayed echolalia. Predominantly, communication is through facial expression and body language. Displays tantrum behaviours which have increased over the last 6 months. |
| | Parents have previously completed HANEN: More than Words training, and try to implement strategies at home to help Meg, but acknowledge time limitations with the demands of Meg and her two younger twin siblings. |
| | Priorities for Meg's parents include building Meg's functional communication skills, decreasing tantrum behaviour at home and having Meg supported as she transitions to school. |
| | (Note: as part of supporting Meg, the SLP may also collect information from teachers and make specific recommendations for support at school; however this will occur later and is not part of the assessment plan described below in this document). |

| Description of Assessment |
|---|
| **Aim of Assessment** |
| Identify goals for a 12-week block of intervention targeting at increasing Meg's range of communicative functions and decreasing tantrum behaviours at home when Meg's communication is not able to be understood.

Describe Meg's current communication abilities in a report for her school in order to assist Meg's teachers to understand her communication abilities and respond appropriately to her communicative behaviours. |
| **Assessment Procedure, Materials & Content** |
| The SLP conducts a face-face interview to obtain information from Meg's mother. The interview is conducted during a clinic appointment, while Meg is at school and her younger siblings are at Day-care. The 0-4 year old questions from the "Pragmatics Profile for Everyday Communication Skills in Children" (Dewart & Summers, 1995) are used to guide the informal interview. Questions cover four areas: Communicative functions; Response to communication from others; Interaction and conversation with others; and variation in communication depending on context. |

The SLP adjusts the questions from the "Pragmatics Profile for Everyday Communication Skills in Children" in order to obtain information relevant to Meg (e.g. Question 2b "How does (child's name) let you know if (he/she) wants to be picked up?"; "picked-up" is changed to "hugged", to be more relevant to

Meg's age and preferences). The SLP also asks further questions in the interview to probe for more information and/or obtain descriptions of specific examples of Meg's communication at home.

Some of these questions include:

**Communicative Functions:**

**Question 1b** (Attention directing to events, objects, other people)
If you and Meg were going along the street or walking in a park and she saw something interesting, what would Meg be likely to do?
(e.g. point, point and vocalize, point and turn to look at you, say a word such as 'look', 'plane')

**Question 2a** (Request for an object)
When you are in the kitchen at home and Meg sees something she wants to eat that is out of reach, how would she let you know?
(e.g. by crying; by reaching out; by pointing and making pleading noises; by pulling you over to it; pointing at the object and saying its name etc).

**Question 2b** (Request for an action)
How does Meg let you know if she wants to be hugged or cuddled?

**Question 2d** (Request for Recurrence)
When you were pushing Meg on the swing at home and she wanted you to do it again, how would she let you know?

**Question 3** (Rejecting)
When Meg is at the table and you are giving her food that she doesn't want, what is she likely to do?

**Question 6** (Naming)
When Meg identifies something she recognises, how does she give it a name?

**Question 7a** (Commenting on Objects)
If you are putting things away and Meg sees something she is interested in, what might she do?

**Question 7b**. (Commenting on Disappearance)
When Meg notices that something at home has gone from where she would usually expect it to be, what sort of comment would she make?

**Response to Communication:**

**Question 10** (Interest in Interaction)
When you are sitting close to Meg and talking to her, how does she generally respond?
(e.g. shows little interest, looks and makes eye-contact, moves body or face)

**Question 16a** (Response to 'No' and Negotiation)
If you have to say 'no' to Meg how does she usually respond?
(e.g. accepts it, has a tantrum, keeps on asking)

**Interaction and Conversation:**
**Question 17** (Participating in Interaction)
When you and Meg are playing or interacting together, how does she take part?

**Question 18** (Initiating Interaction)
If Meg ever starts up a little game with you, how does she do it?
(e.g. by catching your eye, by making little sounds, by coming close to you and looking into your face, by giving or showing you something).

**Contextual variation:**

**Question 28** (Time)
Are there times of the day when Meg is more communicative at home?
(e.g. mealtimes, bath time, at the playground, in the morning)

**Question 30** (Books as a context for communication)
How does Meg respond when you read books to her at home?

**Question 31** (Use of language in play)
When Meg is playing by herself at home, what does she play and what kind of communication goes on?
(e.g. what kind of sounds/gestures does she make?)

---

**Assessment analysis**

---

SLP will use information from the interview to report on Meg's current communication strengths and weaknesses across different communicative functions and identify immediate priorities for communication goals. After this assessment is completed, Meg will receive a block of intervention.

Note: These case studies were created specifically for the purposes of this Delphi Study. They are not intended to be examples of "recommended practice" nor are they intended to represent how assessments are most frequently used in SLP practice.

Reference:
Dewart, H. & Summers, S. (1995). The Pragmatics Profile of Everyday Communication Skills in Children. Available from: www.complexneeds.org.uk/modules/Module-2.4-Assessment-monitoring-and-evaluation/All/downloads/m08p080c/the_pragmatics_profile.pdf.

**Supplementary Appendix 3.7.**

Case Study Two (Assessment Plan)

| Background Information | |
|---|---|
| **Student's name** | Eric |
| **Student's Age** | 7;10 years |
| **Background information** | Eric is in Grade 2 (3rd year of full-time school). Parents report nothing significant in developmental or medical history. He has a history of educational concerns at school. His current teacher reports: immature grammar and sentence structure when talking; often requires support to elaborate on information e.g. during show and tell; difficulties following directions and staying on task in class activities; behind expected level for reading, particularly reading comprehension.<br><br>Eric participated in small group programs in grade 1 targeting oral language and reading, run by school learning support teacher. He currently receives small group support at school for reading, also run by learning support teacher.<br><br>In a recent assessment session at school, the Clinical Evaluation of Language Fundamentals – 4th Edition (CELF-4) was administered. In this current assessment session, language sampling is to be conducted. |

| Description of the Assessment |
|---|
| **Aims of Assessment** |
| Identify practical strategies/supports that his teacher and parents may find useful for facilitating Eric's oral language production in class and at home.<br><br>Provide qualitative descriptions of Eric's difficulties (as further evidence of the difficulties identified on the CELF-4 and in teacher reports).<br><br>Collect a measure Eric's current skills to monitor development over time. |
| **Assessment procedure, materials and content** |
| The SLP will conduct a structured assessment with Eric in a withdrawal room away from regular class activities. The SLP uses the administration procedure and materials in the "Language Sampling Protocol" (Westerveld & Gillon, 2002); with additional prompts only supplied after tasks have been initially completed according the administration procedure.<br><br>**Conversation:**<br>Eric is asked to bring some of his artwork from class to the assessment session. The SLP engages Eric in conversation by asking *"What did you bring to show me?"* If needed to stimulate conversation (and collect at least 50 utterances) the SLP may make comments or ask open-ended questions. The SLP notes the level support Eric needs and the types of supports that assist him in conversation.<br><br>**Narrative comprehension:**<br>Eric listens to the story "Ana gets Lost" once. The SLP then immediately asks him the eight comprehension questions about the story. |

Four questions relate to information directly stated in the story e.g.
*"Who is the story about?"*
*"What did the policeman do*?"
Four questions require inferencing and understanding of causal relations e.g.
*"Why did Anna have to stay at home?"*
*" Why were Anna's parents happy to see her?"*
After all the questions have been asked and Eric's answers recorded, the SLP discusses the answers to any questions that Eric did not know (this is to reduce the impact of poor comprehension on the completion of the next task). During the discussion, the SLP provides additional prompting to observe the level of scaffolding Eric requires to answer the questions.

**Narrative Retell:**
After a short break (in which the SLP completes the conversational sample above), Eric then listens to the story again with the SLP instructing: "*Let's listen to the story a second time. Afterwards we will put a new tape in the recorder and then I would like you to tell the story, so that other children can listen to it later."*
Eric retells the story without pictures.
If required, the SLP prompts with general open-ended questions to help Eric begin his retell e.g.
*"What was the story about?"*
*"What happened in the beginning?"*
*"Just tell me what you remember".*
If required, the SLP may prompt to elicit further information once Eric has begun retelling e.g. *"And then?"* or "*Anything else you can remember?"*
Eric then listens to the story a third time with the SLP instructing: *"Let's try that one more time? You can look at the pictures this time when you're telling the story. Let's start at the beginning."*
Eric retells the story using the picture cues from the book. If required, the SLP may provide prompts (as above).

---

**Assessment Analysis**

---

**Conversational sample:**
The SLP makes a note of the level of support (i.e. number and types of prompts) that Eric needs to participate in conversation with the SLP and the influence of different prompts on his performance (i.e. closed vs open ended questions, additional time etc). The conversational sample recording is also analysed and described with regards to verbal fluency (e.g. mazing) and the types of sentence structures used.

**Narrative comprehension:**
The SLP initially notes Eric's performance on the comprehension questions without support and differences in his performance on the factual versus inferential questions. The SLP then notes Eric's performance when additional scaffolding was provided.

**Narrative retell:**
Eric's performance on the narrative comprehension recording is analysed and described with regards to verbal fluency (e.g. mazing), types of sentence structures used and the narrative structure quality (story grammar). The SLP also compares the quality of the retell with no-picture support versus picture support and a third reading of the story. The SLP also notes Eric's response to different types of prompts.

Note: These case studies were created specifically for the purposes of this Delphi Study. They are not intended to be examples of "recommended practice" nor are they intended to represent how assessments are most frequently used in SLP practice.

Reference:
Westerveld, M. & Gillon, G. (2002). Language Sampling Protocol. Available from:
www.education.canterbury.ac.nz/documents/gillon/languageprotocol.pdf

# Chapter 4.

# Factors Influencing SLPs' Application of a Taxonomy with

# Terminology for Describing Language Assessments

**Overview for Chapter 4 (Journal Article 3)**

Chapter 4 relates to research area two. This chapter further builds upon Chapter 3 by collecting qualitative information to support the implementation of the taxonomy. Semi-structured interviews were used to investigate perspectives of previous Delphi study participants regarding factors that may influence consistent application of the taxonomy terminology by SLPs. Participants also identified strategies to support SLPs use of the taxonomy. These strategies were subsequently incorporated into the design of the survey presented in Chapters 5 and 6. This chapter contains a manuscript currently under review with the following journal: Communication Disorders Quarterly.

**Factors Influencing Speech-Language Pathologists' Application of Terminology for Describing Child Language Assessments**

Deborah Denman[1*], Nathan J. Wilson[2], Natalie Munro [3,1], Jae-Hyun Kim[4, 1,] Renée Speyer [5, 6,] [1], Reinie Cordier[1, 5]


[1] School of Occupational Therapy, Social Work and Speech Pathology, Faculty of Health Sciences, Curtin University, Perth, Australia

[2] School of Nursing and Midwifery, Western Sydney University

[3] Faculty of Health Sciences, The University of Sydney, Sydney, Australia

[4] Department of Linguistics, Macquarie University, Sydney, Australia

[5] Department Special Needs Education, University of Oslo, Oslo, Norway

[6] Department of Otorhinolaryngology and Head and Neck Surgery, Leiden University Medical Centre, Leiden, The Netherlands


Short title: SLP Application of Terminology


*Corresponding author: E-mail: deborah.denman@postgrad.curtin.edu.au.

## 4.1. Abstract

Objective: This study investigated SLPs' perceptions of factors that influence application of a new taxonomy with terminology for describing child language assessment and identified strategies that may facilitate use of taxonomy terminology to collect data on SLP assessment practice.

Method: Semi-structured interviews were conducted with 13 SLPs and data were analysed using thematic analysis.

Results: Three main themes were identified in relation to factors that may influence application including: applying the taxonomy is arduous, contextual factors may influence application, and SLP experience and knowledge may influence application. Participants identified a number of strategies to facilitate use of taxonomy by SLPs.

Conclusion: Findings from this study give insight into the factors that influence SLPs application of a taxonomy of assessment terms. These findings are important for all SLPs in the child language field to consider if the profession is to be effective in establishing greater consistency in use of professional terminology.

**4.2. Introduction**

Lack of consistent terminology is widely acknowledged across the speech-language pathology (SLP) profession (Roulstone, 2015; Walsh & IGOTF-CSD., 2006). Terms to describe clinical practices may be used ambiguously, different terms may be used interchangeably for a single concept or the same term may be used with different interpretations (Cowie et al., 2001; Walsh, 2005). This was highlighted in a previous study by Cowie et al. (2001) examining the terminology used by SLPs in clinical case notes. It was identified that terms used by SLPs were used inconsistently or ambiguously, potentially impacting on the accuracy with which other SLPs would be able to interpret the case notes. Furthermore, it was found that professional terminology was used inconsistently not only between different SLPs, but also between different case notes kept by the same SLP. Ambiguous descriptions of clinical procedures in case notes may have medico-legal consequences as clinical decisions may not be transparent (Cameron & Turtle-Song, 2002).

In another study, SLPs were asked to identify the 'non-standardised' or 'informal' assessment procedures they used to assess children (Roulstone et al., 2015). A range of procedures were listed by SLPs including 'observation', 'play', 'audio-recording', 'language sampling in context', 'picture description' and 'posting games'. The use of such non-specific terms makes it difficult to collect detailed data on clinical practice and creates barriers when attempting to compare current practice to research evidence (Cowie et al., 2001; Eadie, 2003).

Lack of detailed terminology may also hinder reflective thinking, which requires purposeful and critical analysis of one's clinical practice (Caty, Kinsella, & Doyle, 2015; Mann, Gordon, & MacLeod, 2009). Without detailed terms for conceptualizing clinical procedures, it is difficult for SLPs to define and compare the specific features of different procedures and thus fully engage in the reflective thinking practices that are needed for

professional development (Eadie, 2003). Therefore, establishing consistent professional terminology is an important step in advancing clinical practice (Cowie et al., 2001).

To address the need for well-defined terminology in the field of child language, a taxonomy for describing different types of language assessments was recently developed in a Delphi study using an online survey involving 55 Australian paediatric SLPs (Denman, Kim, Munro, Speyer, & Cordier, 2019). The taxonomy was developed with the aim of establishing agreed-upon terminology to facilitate consistency between SLPs with regards to descriptions of language assessment practices. As such, the taxonomy can be used in workplaces to support effective professional communications between SLPs and is particularly useful in situations where assessments need to be described consistently, for example, clinical record keeping, assessment reports, audits or survey research. Although agreement was established using a group of Australian SLPs, the taxonomy was based upon literature from the United States of America and United Kingdom, and thus has applicability across other English-speaking countries (Denman et al., 2019).

The taxonomy was structured around four main aspects by which assessments may be described. Each aspect contains distinct categories with a detailed definition provided for each term in each category. Consensus from SLPs regarding the taxonomy structure, definitions and terms was established during the previous Delphi study, with 100% of participants expressing agreement or strong agreement or with the overall taxonomy structure and at least 88% of participants expressing agreement or strong agreement with the definitions within each aspect (Denman et al., 2019). The taxonomy aspects and categories are presented schematically in Figure 4.1.

**Figure 1. Structure of the taxonomy (Based on Denman et al., 2019)**
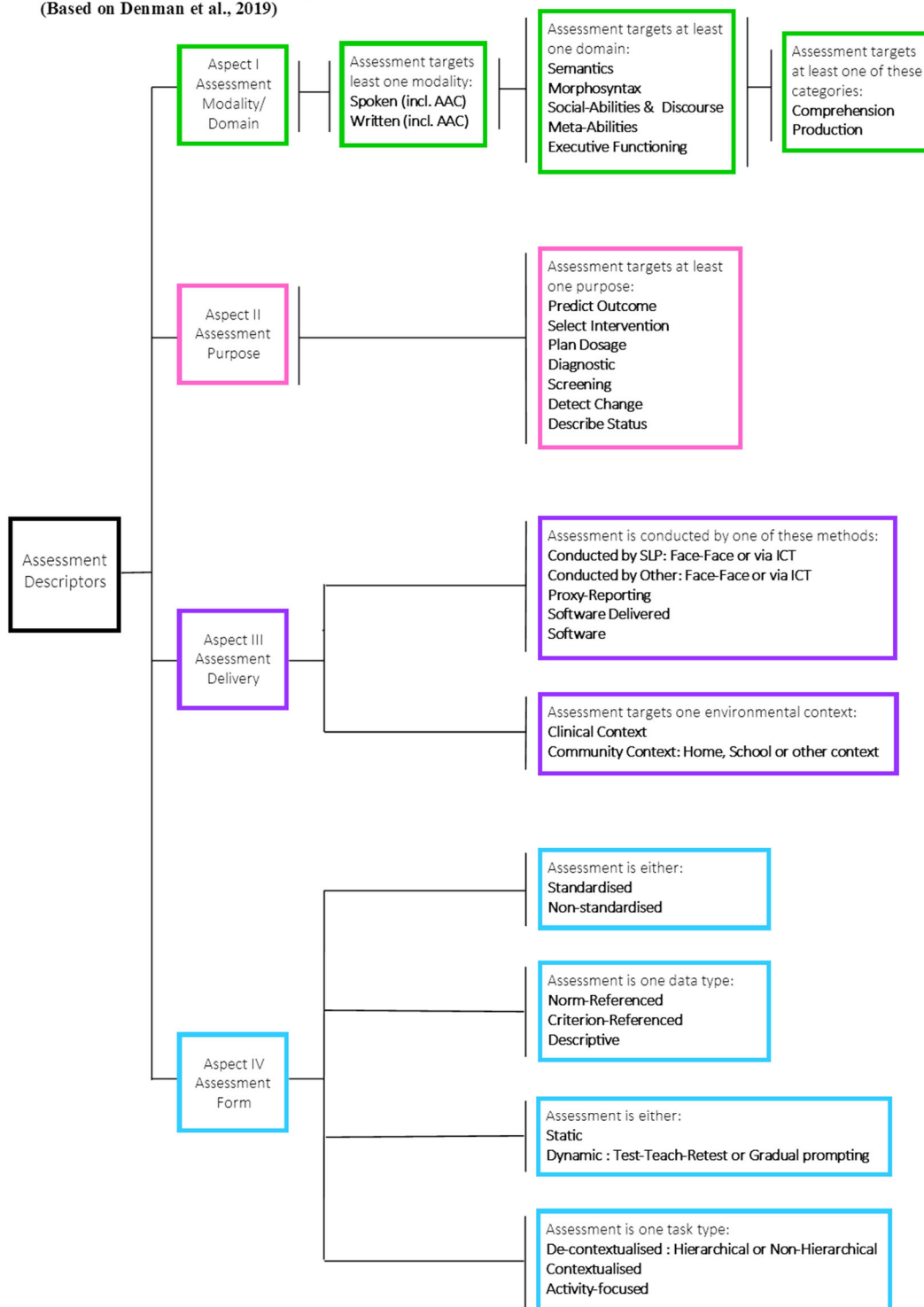
*Figure 4.1.* Taxonomy structure. *Note*: A different version of this same taxonomy is presented in Figure 3.1.

Despite reaching strong agreement on the structure and definitions of the taxonomy, SLPs in the Delphi study were not always consistent when applying the taxonomy to describe assessments presented in case studies. Literature identifies that, although SLPs value professional innovations, they often experience challenges when applying new innovations to clinical practice (Cheung, Trembath, Arciuli, & Togher, 2013; O'Connor & Pettigrew, 2009; Zipoli & Kennedy, 2005). It also acknowledged that implementation of new knowledge is typically not spontaneous and that a variety of factors may influence the successful transfer of new knowledge into everyday practice contexts (Graham et al., 2006; Harding, 2014). Although this newly developed taxonomy does not aim to alter clinical assessment practice itself, adopting the taxonomy may require SLPs to change the terminology that they routinely use to describe language assessments or the current structure that they use to conceptualise different types of language assessments; which may be viewed as a process of knowledge to action transfer (Denman et al., 2019; Harding, 2014; Miao, Power, & O'Halloran, 2015).

Knowledge to action transfer processes are widely discussed across implementation science literature (Graham et al., 2006), however only a small number of previous studies have explored implementation of research knowledge by SLPs (Cheung et al., 2013; Miao et al., 2015; O'Connor & Pettigrew, 2009; Shrubsole, Worrall, Power, & O'Connor, 2018; Young, Shrubsole, Worrall, & Power, 2018; Zipoli & Kennedy, 2005). These previous studies have focussed on implementation of evidence-based research practices more generally (Cheung et al., 2013; O'Connor & Pettigrew, 2009; Zipoli & Kennedy, 2005) or implementation of clinical practice guidelines by SLPs working specifically the area of adult rehabilitation (Miao et al., 2015; Shrubsole et al., 2018; Young et al., 2018). Although consistent application of professional terminology has been a focus in recent SLP literature, particularly in relation to developing consensus on diagnostic terminology (Bishop, 2017; Reilly et al., 2014); no studies exploring the factors that influence SLPs application of

agreed-upon terminologies or taxonomies have not been identified. Therefore, a greater understanding of the challenges associated with implementing consistent terminology needs to be developed. This knowledge will not only facilitate the successful future application of this new taxonomy but will also provide greater understanding of the challenges related to the establishment of consistent terminology across the SLP profession more broadly.

**4.2.1. The current study**. The specific objectives of this study were:

1. To identify SLPs' perceptions regarding factors that may influence consistent application of a taxonomy with terminology for describing child language assessments.

2. To identify strategies that may support future use of the taxonomy by SLPs when describing language assessment practices.

Information gained from this study will assist SLP training providers, clinicians, researchers and service managers with implementing consistent terminology to facilitate professional communication, reflective thinking or accurate data collection on SLP assessment practice.

## 4.3. Method

**4.3.1. Approach**. As this was an exploratory study seeking the views of individual SLPs, this study adopted a qualitative approach using individual semi-structured interviews (McIntosh & Morse, 2015) and thematic analysis of interview data (Braun & Clarke, 2006; Vaismoradi, Hannele, & Bondas, 2013). Interview participants were asked to give their opinions on what may make the taxonomy challenging to apply and what strategies may facilitate consistent use of the taxonomy terminology by SLPs when describing clinical practice. These questions were asked in relation to each of the four taxonomy aspects.

Prior to the start of interview recordings, participants were informed of the background and aims of the study. All participants provided written consent to participate in the interview and for their de-identified responses to be used in data analysis. Ethical

approval for this study was provided by (deleted for peer review). During the study there were no adverse events or participant withdrawals.

**4.3.2. Participants.** Interviews were conducted until data saturation was reached (Elo et al., 2014). Data saturation is obtained when no new themes emerge from consecutive interviews. In this study, the final sample size was 13 participants. This sample size of 13 is also consistent with estimates provided in literature regarding the sample size at which data saturation is reached in interview research (Guest, Bunce, & Johnson, 2006). The demographics of the 13 participants are reported in Table 4.1. To ensure preservation of participant's anonymity, demographic data are presented as a group aggregate.

As perceptions from SLPs who had experience applying the taxonomy were needed, the SLPs who previously completed the Delphi study were ideally suited as participants in this study, as these SLPs had both in-depth knowledge of the taxonomy and experience applying the taxonomy. For this reason, criteria for participation were the same as the criteria for the previous Delphi study and included: 1) eligibility for practicing membership with Speech Pathology Australia, and 2) having spent at least 5 (full-time equivalent) of the last 10 years engaged in professional activities where at least 50% or more of the time related to children aged 4-18 years with a language disorder (includes any children who required support for oral or written language, regardless of aetiology, severity, primary diagnosis or associated co-morbidities). In addition, participants needed to have completed all three rounds of the previous Delphi study and supplied an email contact during completion of Delphi study rounds in order to be contacted for interviews.

Table 4.1

*Demographics of Study Participants*

| Category | Subcategory | Number | Proportion of Delphi Study participants |
|---|---|---|---|
| State | NSW | 2 | 2/6 (33.3%) |
| | QLD | 2 | 2/5 (40.0%) |
| | SA | 4 | 4/6 (66.7%) |
| | TAS | 1 | 1/2 (50.0%) |
| | VIC | 2 | 2/9 (22.2%) |
| | WA | 2 | 2/3 (66.7%) |
| | ACT | 0 | 0/0 (0.0%) |
| | NT | 1 | 0/1 (0.0%) |
| | Total | 13 | 13/32 (40.6%) |
| Current Employment | School Sector (government or non-government) | 8 | 8/16 (50%) |
| | University | 3 | 3/7 (42.9%) |
| | Both University and Private Practice | 2 | 2/3 (66.6%) |
| | Other sectors (not any of above) | 6 | 0/6 (0.0%) |
| | Total | 13 | 13/32 (40.6%) |
| Current role | Clinician | 6 | N/A |
| | Service manager | 2 | N/A |
| | Academic | 2 | N/A |
| | Both Clinician and Academic | 3 | N/A |
| | Total | 13 | N/A |
| Highest qualification (in addition to Bachelor's degree or equivalent) | Diploma or Graduate certificate/s in fields related to teaching or psychology | 3 | N/A |
| | Coursework Masters (Speech Pathology or Education) | 2 | N/A |
| | PhD | 4 | N/A |
| | Currently a MPhil or PhD student | 3 | N/A |
| | Nil additional qualifications | 1 | N/A |
| | Total | 13 | N/A |

| Years of experience | 5-10 years | 2 | 2/5 (40.0%) |
|---|---|---|---|
| | 11-15 years | 4 | 4/8 (50.0%) |
| | 16-21 years | 1 | 1/9 (11.1%) |
| | 21+ years | 6 | 6/10 (60.0%) |
| | Total | 13 | 13/32 (40.6%) |

Note: NSW= New South Wales; QLD= Queensland; SA= South Australia; TAS= Tasmania; VIC= Victoria; WA= Western Australia; *Proportion of the 32 Delphi study participants in each demographic group who completed round three of the previous Delphi study; NA= Data was not collected in relation to these same categories in the Delphi study.

A combination of purposive sampling and convenience sampling was used in this study, with effort made to ensure inclusion of participants from different geographical locations and work sectors (Elo et al., 2014). Participants may have known the researchers through professional or collegial networks but did not have personal connections with members of the research team.

**4.3.3. Data collection and analysis.** The semi-structured interviews were conducted by the first author who is a qualified SLP with training in qualitative research. This author received on-going support from the other members of research team, who all have extensive experience in conducting qualitative research. Interviews were conducted in November and December 2017 and each interview took between 30-40 minutes to complete. Where possible, interviews were conducted via web-conferencing software using either Skype or Zoom; however, six participants without access to cloud-based web-conferencing in their workplaces were interviewed via telephone. Each participant was interviewed only once, except in the case of one participant where a technical problem required the interview to be stopped and completed on a following day. Aside from the interviewer and the participant, no other people were present during the actual interview.

A pre-established semi-structured protocol with open-ended questions was used to guide interviews. A copy of the interview questions is provided in Supplementary Appendix

4.1. To assist participant's recall of the taxonomy, participants were provided with a copy of the taxonomy and the case studies to review before the interview and these documents were also available during the interview for participants to refer to. During the interviews, prompting was provided to clarify responses or extract further information from participants if required. Interviews ended when all questions were answered and participants indicated that they had no further information to add.

An online audio to text service was used to create written transcriptions of all interviews. These transcripts were then checked for reliability against the original audio-recordings and interview notes by the first author. Participants were also sent copies of transcripts to double-check. Using an inductive approach, data were analysed by the first author according to the six phases of thematic analysis outlined by Braun and Clarke (2006). In phase one, familiarisation with the data occurred by listening to the audio recordings and reading the transcripts multiple times to identify features and patterns. In phase two involved organising the data into groups based on underlying interpretation of meaning. A latent level of interpretation was used to identify factors that may influence application of the taxonomy and a manifest level of interpretation was used to identify strategies that may facilitate application of the taxonomy (Vaismoradi et al., 2013). In phase three, groups were sorted into themes and themes represented schematically. In phase four the research team critically reviewed the themes to ensure that themes were meaningful in representing the data (Vaismoradi et al., 2013). Phase five involved defining each theme and generating a descriptive name for each. In phase six, the authors discussed the documentation of themes and selected quotes to represent each theme.

**4.3.4. Rigour.** Several steps were undertaken to ensure rigour within this study (Elo et al., 2014; Thomas & Magilvy, 2011). Confirmability was addressed through the use of semi-structured interviews, which allowed the researcher to follow the direction of the interviewee,

rather than leading the interviews (Thomas & Magilvy, 2011). To assist in maintaining objectivity, the interviewer recorded field notes after each interview and also listened back to the audio recordings from each interview and recorded reflections regarding effectiveness of interview techniques, initial thoughts regarding themes, and any potential sources of bias (Thomas & Magilvy, 2011). To limit the influence of bias the interview protocol was developed with input from all authors and was reviewed after completion of the first interview. During development, interview questions were checked for clarity of wording or ambiguous meanings. To assist with transferability, participants from different agencies and geographical locations were selected with participant demographics reported in the results (Elo et al., 2014). To assist with establishing dependability of findings, analysis involved reviewing relevancy of identified themes across different transcripts with an audit trail created (Elo et al., 2014). A detailed description of themes is reported with examples of participant quotes that relate to each theme (Braun & Clarke, 2006). To enhance credibility, peer debriefing occurred throughout the data analysis process with final themes discussed and agreed-upon by the research team (Elo et al., 2014; Vaismoradi et al., 2013).

**4.4. Results**

  **4.4.1. Factors that influence application of the taxonomy.** Three core themes relating to factors that may influence application of the taxonomy were identified in this study. These themes include: 1) applying the taxonomy is arduous, 2) contextual factors may influence application of the taxonomy, and 3) SLP experience and knowledge may influence application of the taxonomy. These three core themes encompass 12 subthemes as illustrated in Figure 4.2.

**Factors that influence consistent application of taxonomy terminology**

**SLP experience and knowledge**

1. SLPs may tend to think about the language measure itself rather than the situation in which the measure is used

2. Some concepts may not be well understood generally within the profession

3. Confusion with some taxonomy definitions

4. SLPs are not accustomed to reflecting on assessment to the depth the taxonomy requires

5. Distinctions between taxonomy categories are less definitive in clinical practice

6. SLP values regarding consistent terminology or reflective practice may vary

**Contextual factors**

1. SLP experience or perspectives may influence application

2. Familiarity with assessments used in case studies may influence application

3. Workplace factors may influence application

**Applying the taxonomy is arduous**

1. Need to read taxonomy definitions carefully as terms are not always intuitive

2. Lots of information to read and process when applying the taxonomy

3. It takes time to develop depth of understanding with the taxonomy

*Figure 4.2.* Themes regarding factors that influence consistent application of the taxonomy

***4.4.1.1. Theme one: Applying the taxonomy is arduous.*** This theme relates to SLPs needing time to become familiar with the taxonomy. All participants commented that the taxonomy category names were not always intuitive and may be miss-applied if SLPs do not take the time to read the taxonomy definitions carefully. These comments formed the first subtheme within this theme 'applying the taxonomy is arduous'. For example, participant 07 commented:

> *...So, I think the detail is there and it does explain it, but I think again, if people just look at the headings [category names], they may misunderstand that, or misinterpret that.* (Participant 07)

This subtheme of category names not being intuitive was identified with regards to all aspects of the taxonomy but was a particularly significant subtheme for the *environmental context* categories (see Figure 4.1; Aspect III). Specifically, participants commented on the potential for *environmental context* to be mis-interpreted as the physical location in which the assessment takes place.

Most participants also commented on the cognitive load associated with amount of reading that was required during the Delphi study in order to understand the structure and distinctions between categories in the new taxonomy. These comments related to the second subtheme within 'applying the taxonomy is arduous'. For example, participant 01 commented:

> *...it does feel quite labour-intensive, I suppose, reading the chunks of information, the paragraphs of information; and trying to sort them out in my own head, as to what each aspect is referring to. Because it is quite brain fatiguing, I suppose, to work all those things out.* (Participant 01)

SLPs also identified that it took time to develop the depth of understanding required to apply the taxonomy or that applying the taxonomy became easier with repeated exposure over Delphi rounds. These comments comprised the third subtheme related to 'applying the taxonomy is arduous'. For example, participant 04 commented:

*You know - it potentially becomes a bit of a growing concept about being so explicit and clear about those different aspects of the assessment form. I don't know, maybe that just going to be a time thing... I don't know whether you could go straight like, 'OK tomorrow, this is what you have to use', I think it would be potentially a growth thing, rather than a straight away thing.* (Participant 04)

***4.4.1.2. Theme two: Contextual factors may influence application.*** This theme describes contextual factors that may influence application of the taxonomy. Some participants commented that different professional viewpoints or perspectives may influence application of the taxonomy, particularly in relation to Aspects I (Domain) and II (Purpose). These comments created the first subtheme within 'contextual factors'. For example, participant 11 commented:

*I think it might come back to your view about language. But a lot of people think of semantic/pragmatic together. I think quite a few people would probably think that semantics is inherent in any type of skill [that is being assessed].* (Participant 11)

SLPs also commented that the degree of familiarity with the particular assessments used in the case studies may be an influencing factor. While this may be less of a consideration in contexts where SLPs are describing their own practice, rather than case studies; it does identify more broadly that the degree of familiarity with the assessments that are being described may be a factor that influences application. These comments related to the second subtheme within 'contextual factors'. For example, participant 01 commented:

*...I wasn't familiar with the assessment and - yeah, it sounded like the*

*assessment came with its own ...But I guess if speech pathologists are reporting on*

*what they do as opposed to trying to understand a case study, then it might make it a*

*bit easier...* (Participant 01)

In addition, two participants who worked in a government (public) education department commented that some prognostic purposes (see Figure 4.1; Aspect II), such as *selecting intervention* and *planning dosage*, are influenced by workplace policy and caseload constraints, more so than by assessment findings. These comments formed the third subtheme within 'contextual factors'. For example, participant 04 commented:

*...and the dosage, again, if you're in private [practice] you can plan that, or in health*

*[agencies] you could probably do agreed sets of intervention; but, for us, again, it's*

*what we can negotiate with the school but sometimes we can't have a huge amount of*

*power over how much actual intervention the child gets.* (Participant 04)

***4.4.1.3. Theme three: SLP professional experience and knowledge may influence***

***application.*** This theme related to comments regarding SLPs understanding of how to apply the taxonomy, which may stem from broader issues relating to professional knowledge. It was identified that SLPs may tend to focus on describing the particular language measure being used, rather than describing how the measure was used in a case study. These comments related to the first subtheme within 'SLP professional experience and knowledge'. For example, participant 06 commented:

*...the issue there would be, is clinicians thinking from a very practical perspective,*

*going, 'Well, yeah, I would actually potentially use this assessment for a hundred*

*different purposes.* (Participant 06)

Participants also commented that some concepts may not be well understood generally across the profession, thus making them more difficult for SLPs to identify and describe

consistently, even in the presence of terminology. Such concepts identified by participants included: executive functioning (See Figure 4.1; Aspect I), meta-abilities (See Figure 4.1; Aspect I), prognostic purposes such as planning dosage (See Figure 4.1; Aspect II), screening assessment (See Figure 4.1; Aspect II), dynamic assessment (See Figure 4.1; Aspect IV) and standardised assessment (See Figure 4.1; Aspect IV). These comments formed a second subtheme within 'SLP professional experience and knowledge'. For example, participant 05 commented:

> *I think, again, people are really unclear about why you might use dynamic assessment and what information that's going to give you, if that makes sense. Like, it's a lack of clarity in the field about that as a means of assessment.* (Participant 05)

Participants also identified possible causes of confusion with some taxonomy definitions. This confusion predominantly related the categories that participants raised as concepts that may not be well understood across the profession in general. This was a third subtheme within 'SLP professional experience and knowledge'. For example, participant 01 commented:

> *So, if I was to administer CELF [Clinical Evaluations of Language Fundamentals] as like a baseline, and then do some work and then retest; would that classify as dynamic [test-teach-retest]? Not static, even though I'm using a static? I don't even know.* (Participant 01)

Many participants made comments related to SLPs not being accustomed to reflecting on assessment practices to the degree that the taxonomy required. These comments related to a fourth subtheme within 'SLP professional experience and knowledge'. For example, participant 03 commented:

*I think speech pathologists are not very good at really thinking through what the*

*purpose of the assessment is, actually. From my experience they just do one.*

(Participant 03)

Similar to this, participants identified that SLPs may have difficulty transferring theoretical knowledge into clinical practice. In particular, participants commented that the categories in Aspects I (Modalities and Domain) and II (Assessment Purpose) may not seem as definitive in clinical practice, as they may seem in theory. This was a fifth subtheme within 'SLP professional experience and knowledge may influence application'. For example, participant 05 commented:

*I think we're often very integrated in our thinking, so when we ask people to pull it*

*(apart) and tease it apart, it's sometimes difficult because it's so integrated. One thing*

*depends on another and they inter-relate, so how people interpret that is really*

*different, I guess - is my experience anyway.* (Participant 05)

It was also identified that application of the taxonomy may be influenced by participants' views on the taxonomy itself and their motivation to apply the taxonomy. This comprised the final subtheme within 'SLP professional experience and knowledge'. During interviews, the majority of study participants expressed value in having detailed and consistent terminology for describing different types of language assessment practices. For example, participant 04 commented:

*I look at education or medicine on the other side and they have such - have much*

*stronger theoretical underpinnings for a lot of their stuff than I think we do. I think it*

*[the taxonomy] is really good; and then, if we get to the stage of embedding this sort*

*of stuff, the sort of conversations that we can have with people can be so much more*

*explicit.* (Participant 04)

In contrast, two study participants presented a different view. These participants did not appear to identify value in having consistent terminology to support reflective thinking, which may influence motivation with regards to taking the time to apply the taxonomy. For example, participant 13 commented:

*...I start to wonder what the point of it is. I wonder if other people would start to question why they are spending time rating to such a detailed, complex level of description. ... So, I think the more people understand what the purpose is for doing something, the more people might engage with it.* (Participant 13)

**4.4.2. Strategies that may facilitate use of the taxonomy.** Study participants identified specific strategies that may support future use of the taxonomy. These strategies fall into two categories: strategies to develop SLPs understanding of the taxonomy and strategies to facilitate use of the taxonomy to collect data on SLP assessment practice. These strategies are summarised in Table 4.2.

**4.5. Discussion**

This study investigated SLP perceptions of factors that may influence application of a taxonomy for describing different types of child language assessments. The finding that numerous factors influence application of the taxonomy is consistent with the outcomes of previous studies investigating implementation of evidence-based practice recommendations by SLPs (Cheung et al., 2013; Miao et al., 2015). These factors identify that even when terminology is agreed-upon, additional actions may be needed to facilitate consistent use of terminology.

SLPs in this study identified strategies to facilitate SLPs' understanding of the taxonomy. These strategies are useful for those who are developing SLP knowledge of the taxonomy through university training or continuing professional development for the purposes of enhancing professional communication and reflective practice. Strategies to

Table 4.2

*Strategies to Facilitate Use of the Taxonomy*

| Categories | Sub-categories | Example of Participant Quote |
|---|---|---|
| Strategies to facilitate SLP's understanding of the taxonomy | Provide case studies to work through | *I think people [SLPs] would need some training to do that [apply the taxonomy], but I think it would be worth it. I think, to begin with, what would be really good is to use your case studies... (*Participant 12) |
| | Provide video explaining taxonomy | *I don't know if some sort of training or maybe if there's like a little - it's not an induction but even like a five-minute or ten-minute little video that maybe SLPs can watch of you explaining or giving an example. That might be easier. That might be more accessible than having to read four pages.* (Participant 01) |
| | Provide examples to explain definitions | *"...whether or not there were just some examples. For example, this certain assessment was used to show that this child needed X, Y and Z with this type of intervention or something.* (Participant 04) |
| | Provide flowcharts for depicting the structure of the taxonomy | *I really, really like the flowchart and the diagrams and I really like the way that you've broken that down, so that you've got the whole [structure] and then you got like, the structure of Aspect I, that really, really helped me.* (Participant 05) |
| Strategies to support SLP's use of the taxonomy when describing their assessment practices | Pre-categorise assessments for SLPs | *Even, you know, a list of all the common assessments that the SLPs would use and having those already pre-categorized and linked into the taxonomy.* (Participant 01) |
| | Allow SLPs to complete data collection in stages | *I just thought it was good that you could do it [categorise the case studies] in your own time, but it can be done in sections and not all at once.* (Participant 01) |
| | Place category definitions on same page as questions | *Whether it's in a format where there could be boxes that could give those immediate reminders when you need them, rather than flicking back through quite a comprehensive document to find the definitions and descriptions you're looking for.* (Participant 07) |
| | Break decision-making into steps | *I don't know, providing them [SLPs] with a guided step through thing where the information they needed is actually right there next to the questions or something. ...almost like a workbook... I think you have to make it small steps and a very good structure to it...* (Participant 03) |
| | Focus on category definitions rather than category names or taxonomy structure | *I think how you lay it out in the questionnaire, you probably may not actually show them [SLPs] the flowchart and what your thought processes are but actually just go through each of those areas that you're looking at in the taxonomy and ask specific questions about those.* (Participant 11) |
| | Highlight key information in definitions | *...really focusing clinician's [SLPs] attentions on the key parts of what makes something contextualized versus not; is probably going to be the best way to get around it.* (Participant 06) |
| | Simplify by reducing information | *Somehow simplifying and picking and choosing just the key, really, really necessary information to gather, and maybe eliminating the less critical information. ...either very simple definitions or completely transparent label.* (Participant 13) |
| | Have primary and secondary category choices (aspect II) | *...even an option to have a primary purpose and then some secondary or additional purposes. You know, SLPs feel like they can have those things reflected even if it's not the primary: 'Oh yes, it's mainly semantics but we do a little bit of meta-abilities, can I put that somewhere?* (Participant 01) |

support use of the taxonomy for collecting data on SLP assessment practice were also identified. These are useful for those who use the taxonomy terminology for purposes such as clinical record-keeping, audits or survey research.

**4.5.1. Factors that influence application of the taxonomy**. SLPs identified that developing understanding of the taxonomy is a learning process that takes time. This time requirement needs to be acknowledged by individual SLPs and within workplaces, particularly as lack of time has previously been identified as a barrier to application of new knowledge by SLPs (O'Connor & Pettigrew, 2009; Zipoli & Kennedy, 2005). SLPs also identified that application of the taxonomy improved with repeated exposure and practice, which highlights the importance of providing information on the taxonomy to SLPs in a format that allows for on-going learning, as opposed to one-off information sessions (Birman, 2000).

Contextual factors, such as differing professional viewpoints or workplace contexts, were also identified as potential influencing factors in the application of taxonomy Aspects I (Modalities and Domains) and II (Assessment Purpose). This finding highlights the need to consider the potential influence of SLP experience and workplace context when SLPs describe assessments. In addition, some concepts related to child language assessment may not be well understood across the profession and these concepts may continue to be problematic even in the presence of well-explained, agreed upon definitions. This indicates an urgent need for further SLP development in SLP training programs and in continuing professional development. Specifically, focus should be placed on developing professional knowledge of concepts: executive functioning, meta-abilities, dynamic assessment and purposes of assessments.

Attention may also need to be placed on the importance of developing consistent and detailed terminology for supporting professional communication and reflective thinking

practices (Caty et al., 2015; Mann et al., 2009). Implementation of new innovations is more successful when clinicians see the advantage of implementing such innovations (Michie et al., 2005) and have support from peers and managers (Cheung et al., 2013; O'Connor & Pettigrew, 2009). Therefore, the importance of establishing consistent use of professional terminology needs to be highlighted across the profession.

**4.5.2. Strategies for developing SLPs' understanding of the taxonomy.** Strategies that may be used in training to facilitate SLPs' understanding of the taxonomy include the provision of training materials such as video presentations, examples of scenarios depicting different types of assessments, and provision of case studies using different assessments. When taxonomy definitions are different to SLPs' pre-existing knowledge of terms, understanding may be enhanced through training materials that draw a direct comparison of how new taxonomy terms compare with individual SLPs' current understanding of terms (Miao et al., 2015). The use of flowcharts may also help make written information about the taxonomy more accessible to SLPs (Gagliardi, Brouwers, Valerie A Palda, Lemieux-Charles, & Grimshaw, 2011).

**4.5.3. Strategies to support use of the taxonomy for collecting data.** Strategies to support use of the taxonomy by SLPs include actions that reduce attention, memory and other cognitive demands associated with synthesising information (Michie et al., 2005). This includes simplifying information by focusing on the definitions of terms, rather than explaining the taxonomy structure; and providing a step-by-step process for making decisions about terms to use to describe different assessments. It may also be important to direct SLPs to the definitions of categories, rather than the category names themselves, in order to avoid misinterpretations of category names. Pre-categorising assessments using taxonomy terms was also suggested as a strategy to reduce cognitive load, as was ensuring that all the information needed to describe assessments was provided together in one place. Allowing

participants to complete data collection over several sessions may assist by providing a break from intense cognitive processing. Consistent application of the taxonomy may also be enhanced when SLPs are conversant with the assessments they are describing.

**4.5.4. Limitations and future directions.** A limitation of this study is that SLPs who were willing to participate in an interview, in addition to a three round Delphi study, may have specific views on the taxonomy and this may have influenced the outcome. Using SLPs who previous participants in the Delphi study allowed for strategies to support use of taxonomy to be identified prior to further implementation of the taxonomy, however it is acknowledged that this group of SLPs may not be representative of the broader SLP population and that a different group of SLPs may experience different challenges with applying the taxonomy. In addition, although all attempts were made to sample SLPs from various workplaces, it is possible that some groups (such as SLPs working in health agencies) are under-represented and that factors unique to these work contexts may not have been explored. As more SLPs gain experience in using the taxonomy, further research should build on the findings from this study by investigating perceptions from  SLPs with different demographics to the participants in this study.

It is acknowledged that both the semi-structured interviews and data-analysis were conducted by the same researcher and bias may have been introduced due to the researcher's pre-existing knowledge and experiences. Potential bias was reduced as much as possible through the use of a reflective journal and through frequent discussions with the broader research team during data interpretation (Thomas & Magilvy, 2011).

It is also important to note that this study was an exploratory study investigating SLPs perceptions of strategies that may facilitate use of the taxonomy. Future research measuring the outcomes of applying these strategies is necessary to evaluate the effectiveness of these strategies for achieving consistent application (Graham et al., 2006). Given the need to

develop consistent terminology across the SLP profession (Cowie et al., 2001; Walsh, 2005), further research is also needed to examine if similar strategies may be used to enhance SLPs application of other terminologies that are used within the profession.

## 4.6. Conclusion

Given the problems that inconsistent terminology poses, establishing consistent terminology is an important endeavour for the SLP profession. The findings from this study give greater insight into factors that influence consistent application of a new taxonomy with terminology for describing different types of child language assessments. These factors include: factors relating to the application of the taxonomy being arduous, contextual factors such as SLPs' perceptions or work contexts, and factors related SLP professional experience and knowledge. This study also identified strategies that may facilitate SLP understanding of the taxonomy and strategies to support use of the taxonomy to collect data from SLPs on the assessment practices they use. This information provides direction for SLP training providers, clinicians, researchers and service managers when establishing consistent application of terminology for describing child language assessments. In addition, findings from this study provide a greater understanding of the challenges associated with establishing consistent professional terminology more broadly.

**References for Chapter 4**

Birman, B. F. (2000). Designing professional development that works. *Educational Leadership, 57*(8), 28-33.

Bishop, D. V. M. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language and Communication Disorders, 52*(6), 671-680. doi:10.1111/1460-6984.12335

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa

Cameron, S., & Turtle-Song, I. (2002). Learning to write case notes using the SOAP format. *Journal of Counseling & Development, 80*(3), 286-292. doi:10.1002/j.1556-6678.2002.tb00193.x

Caty, M. È., Kinsella, E. A., & Doyle, P. C. (2015). Reflective practice in speech-language pathology: A scoping review. *International Journal of Speech-Language Pathology, 17*(4), 411-420. doi:10.3109/17549507.2014.979870

Cheung, G., Trembath, D., Arciuli, J., & Togher, L. (2013). The impact of workplace factors on evidence-based speech-language pathology practice for children with autism spectrum disorders. *International Journal of Speech-Language Pathology, 15*(4), 396-406. doi:10.3109/17549507.2012.714797

Cowie, M., Wanger, K. M., Cartwright, A., Bailey, H., Millar, J. A., & Price, M. (2001). A review of Clinical Terms Version 3 (Read Codes) for speech and language record keeping. *International Journal of Language & Communication Disorders, 36*(1), 117-126. doi:10.1080/13682820116848

Denman, D., Kim, J.-H., Munro, N., Speyer, R., & Cordier, R. (2019). Describing language assessments for school-aged children: A Delphi study. *International Journal of*

*Speech Language Pathology*, Advance Online Publication.

doi:10.1080/17549507.2018.1552716

Eadie, P. (2003). Speech pathology assessment practices: One assessment or many? *Advances in Speech Language Pathology, 5*(1), 65-68. doi:10.1080/14417040510001669081

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative content analysis: A focus on trustworthiness. *Sage Open, 4*(1). doi:10.1177/2158244014522633

Gagliardi, A. R., Brouwers, M. C., Valerie A Palda, V. A., Lemieux-Charles, L., & Grimshaw, J. M. (2011). How can we improve guideline use? A conceptual framework of implementability. *Implementation Science, 6*, 26. doi:10.1186/1748-5908-6-26

Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J., Caswell, W., & Robinson, N. (2006). Lost in knowledge translation: Time for a map? *The Journal of Continuing Education in the Health Professions, 26*, 13-24. doi:10.1002/chp.47

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59-82. doi:10.1177/1525822X05279903

Harding, K. E. (2014). Not enough time or a low priority? Barriers to evidence-based practice for allied health clinicians. *Journal of Continuing Education in the Health Professions, 34*(4), 224-231. doi:10.1002/chp.21255

Mann, K., Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: a systematic review. *Advances in health sciences education, 14*(4), 595-621. doi:10.1007/s10459-007-9090-2

McIntosh, M. J., & Morse, J. M. (2015). Situating and constructing diversity in semi-structured interviews. *Global Qualitative Nursing Research, 2*, 1-12. doi:10.1177/2333393615597674

Miao, M., Power, E., & O'Halloran, R. (2015). Factors affecting speech pathologists' implementation of stroke management guidelines: A thematic analysis. *Disability and Rehabilitation, 37*(8), 674-685. doi:10.3109/09638288.2014.932444

Michie, S., Johnston, M., Abraham, C., Lawton, R., Parker, D., & Walker, A. (2005). Making psychological theory useful for implementing evidence based practice: A consensus approach. *Quality and Safety in Healthcare, 14*(1), 26-23. doi:10.1136/qshc.2004.011155

O'Connor, S., & Pettigrew, C. M. (2009). The barriers perceived to prevent the successful implementation of evidence-based practice by speech and language therapists. *International Journal of Language & Communication Disorders, 44*(6), 1018-1035. doi:10.3109/13682820802585967

Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., . . . Wake, M. (2014). Terminological debate over language impairment in children: forward movement and sticking points. *International Journal of Language & Communication Disorders, 49*(4), 452–462. doi:10.1111/1460-6984.12111

Roulstone, S. (2015). Exploring the relationship between client perspectives, clinical expertise and research evidence. *International Journal of SpeechLanguage Pathology, 17*(3), 211-222. doi:10.3109/17549507.2015.1016112

Roulstone, S., Marshall, J., Powell, G. G., Goldbart, J., Wren, Y. E., Coad, J., . . . Coad, R. A. (2015). Evidence-based intervention for preschool children with primary speech and language impairments: Child Talk – an exploratory mixed-methods study. *Programme Grants for Applied Research, 3*(5), 1-408. doi:10.3310/pgfar03050

Shrubsole, K., Worrall, L., Power, E., & O'Connor, D. A. (2018). Barriers and facilitators to meeting aphasia guideline recommendations: What factors influence speech pathologists' practice? *Disability and Rehabilitation*, Advance online publication. doi:10.1080/09638288.2018.1432706

Thomas, E., & Magilvy, J. K. (2011). Qualitative rigor or research validity in qualitative research. *Journal for Specialists in Pediatric Nursing, 16*, 151-155. doi:10.1111/j.1744-6155.2011.00283.x

Vaismoradi, M., Hannele, T., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & health sciences, 15*(3), 398-405. doi:10.1111/nhs.12048

Walsh, R. (2005). Meaning and purpose: A conceptual model for speech pathology terminology. *Advances in Speech-Language Pathology, 7*(2), 65-76. doi:10.1080/14417040500125285

Walsh, R., & IGOTF-CSD. (2006). A history of terminology: International group on terminology frameworks – communication science and disorders. . Retrieved from http://www.dhrs.uct.ac.za/sites/default/files/image_tool/images/147/History_of_CSD.pdf

Young, L., Shrubsole, K., Worrall, L., & Power, E. (2018). Factors that influence Australian speech-language pathologists' self-reported uptake of aphasia rehabilitation recommendations from clinical practice guidelines. *Aphasiology, 32*(6), 646-665. doi:10.1080/02687038.2018.1443201

Zipoli, R. P., & Kennedy, M. (2005). Evidence-based practice among speech-language pathologists: Attitudes, utilization, and barriers. *American Journal of Speech-Language Pathology, 14*, 208-220. doi:1058-0360/05/1403-0208

**Supplementary Appendix 4.1.**

Semi-Structured Interview Protocol

**<u>Introduction</u>**

Introductions, explain study and clarify any information about the study if needed

**<u>Overall Question</u>**

**Considering the taxonomy, what are your thoughts about using the taxonomy to survey SLPs on their assessment practices?**

*Follow-up probes (to be used as needed):*

- *What would you suggest doing if the taxonomy was to be used in a survey?*

- *Are there any strategies that you think would facilitate consistent use of the taxonomy?*

- *Why do you think that?*

- *Tell me more about _____?*

**<u>Taxonomy Aspect I</u>**

What about Aspect I of the taxonomy? *(Show reference sheet)* In this aspect, there was consensus on the main domains targeted in the case studies but a lack of consensus on whether categories *semantics* and *executive functioning* were targeted as secondary domains.

**Question 1a): What do you think could make applying aspect I difficult** (if not covered previously)?

*Follow-up probes (to be used as needed):*

- *Why do you think that?*

- *Did you find applying aspect I difficult (and why)?*

- *Why do you think the categories semantics and executive functioning in particular, were applied inconsistently in case studies?*

**Question 1b): What do you think would facilitate using aspect I in a national survey (if participant did not cover this in previous answer)?**

*Follow-up probes (to be used as needed):*

- *Why do you think that would assist?*

**Taxonomy Aspect II**

What about Aspect II of the taxonomy? *(Show reference sheet)* In this aspect, there was lack of consensus on purposes of assessments.

**Question 2a): What do you think could make applying aspect II difficult** (if not covered previously)? *Follow-up probes (to be used as needed):*

- *Why do you think that?*

- *Did you find applying aspect I difficult (and why)?*

**Question 2b): What do you think would facilitate using aspect II in a national survey (if participant did not cover this in previous answer)?**

*Follow-up probes (to be used as needed):*

- *Why do you think that would assist?*

**Taxonomy Aspect III**

What about Aspect II of the taxonomy? *(Show reference sheet)* In this aspect, there was lack the categories for *environmental context* were particularly inconsistently applied.

**Question 3a): What do you think could make applying aspect III difficult** (if not covered previously)? *Follow-up probes (to be used as needed):*

- *Why do you think that?*

- *Did you find applying aspect III difficult (and why)?*

- *Why do you think the categories for environmental context in particular, were applied inconsistently in case studies?*

**Question 3b): What do you think would facilitate using aspect III in a national survey (if participant did not cover this in previous answer)?**

*Follow-up probes (to be used as needed):*

- *Why do you think that would assist?*

**Taxonomy Aspect IV**

What about Aspect II of the taxonomy *(show reference sheet)*. In this aspect, there was lack the categories across a number of components including '*standardised* vs. *non standardised*', '*static*' vs. '*dynamic*' and '*task type*'.

**Question 4a): What do you think could make applying aspect IV difficult** (if not covered previously)? *Follow-up probes (to be used as needed):*

- *Why do you think that?*

- *Did you find applying aspect III difficult (and why)?*

- *Why do you think the categories for standardised vs. non-standardised in particular, were applied inconsistently in case studies?*

- *Why do you think the categories for static vs. dynamic in particular, were applied inconsistently in case studies?*

- *Why do you think the categories for task type in particular, were applied inconsistently in case studies?*

**Question 4b): What do you think would facilitate using aspect IV in a national survey (if participant did not cover this in previous answer)?** *Follow-up probes (to be used as needed):*

- *Why do you think that would assist?*

**Conclusion**

**Question 5: Before we finish, do you have any other comments about the application of the taxonomy that have not already been covered?**

Close interview and thank participant.

**Chapter 5.**

**Language Assessment Practices for Primary School Children (Part I):**

**What Factors Influence Assessment Use?**

**Background to Chapter 5 (Journal Article 4)**

Chapter 5 relates to research area three. This research complements the research presented in the preceding chapters by providing survey data on the language assessment practices used by Australian SLPs to assess primary school children. Chapter 5 presents findings from Part I of the survey which investigated the regularity with which different language assessments are used, factors that influenced regularity of assessment use, the main challenges reported by SLPs in relation to language assessment for school-aged children and the main sources of information from which SLPs reported obtaining information on language assessment practice. The manuscript contained in this chapter is currently under review with the following journal: Journal of Communication Disorders.

# Language Assessment for Primary School Children (Part I): What Factors Influence Assessment Use?

Deborah Denman[1*,] Reinie Cordier[1,4], Jae-Hyun Kim[2], Natalie Munro[3], Renee Speyer[4,5,1]

[1] School of Occupational Therapy, Social Work and Speech Pathology, Faculty of Health Sciences, Curtin University, Perth, Australia

[2] Department of Linguistics, Macquarie University, Sydney, Australia

[3] Faculty of Health Sciences, The University of Sydney, Sydney, Australia

[4] Department Special Needs Education, University of Oslo, Oslo, Norway

[5] Department of Otorhinolaryngology and Head and Neck Surgery, Leiden University Medical Centre, Leiden, The Netherlands

*Corresponding author: E-mail: deborah.denman@postgrad.curtin.edu.au

**5.1. Abstract**

Background: Language assessment is an important component in service provision for children with language disorder, therefore it is important that focus is placed on the types of language assessments that speech-language pathologists (SLP)s use to assess children.

Objective: This study reports on data from Part I of a survey of SLP language assessment practices for primary school children. The objective of the study was to investigate the regularity with which SLPs use different types of assessments when assessing the language abilities of children aged 4-12 years. This study also investigated the factors that influence the regularity with which different types of assessments are used, the challenges reported by SLPs in relation to language assessment and the main sources of information from which SLPs obtain information on assessment practices.

Methods: A web-based survey was used to collect information from 407 Australian SLPs regarding the assessments practices they use for school-aged children. Terms and definitions from a recently developed taxonomy were used to guide the development of survey questions and provide explicit terminology for describing types of language assessments. Factors that influenced the regularity with which different types of assessments were used were investigated using regression analysis.

Results: Most SLPs regularly used assessments that are norm-referenced, de-contextualised, and conducted in a clinical context and less regularly used other types of assessments. Service agency, Australian State, and SLP years of experience were found to influence the regularity with which different types of assessments were used. The main challenges reported by SLPs related to limited time, lack of assessment materials, limited access to training, and lack of confidence in assessing children from culturally and linguistically diverse

backgrounds. Informal discussion with colleagues was the most frequently identified source of information on assessment practice.

Conclusion: SLPs could improve current language assessment practice for primary school children through more regular use of assessments that are contextualised, activity-focused, dynamic or targeted at school, home or community contexts. Actions to facilitate increased use of these assessments should consider the contextual differences that exist between service agencies and states and address challenges that SLPs experience in relation to language assessment.

**5.2. Introduction**

Language disorder is identified when a child has persistent difficulties with spoken and written language with these difficulties impacting on everyday social interactions and educational achievement (Bishop, Snowling, Thompson, Greenhalgh, & CATALISE-2 consortium, 2017). Speech language pathology (SLP) intervention in the school years is important for reducing the functional limitations that children with language disorder experience (Norbury et al., 2016). As interventions and supports for children are determined based on assessment data, it is important that attention is placed on the types of assessments that SLPs use when assessing the language abilities of primary school children (Eadie, 2003).

A wide variety of assessment options exist for assessing children with language disorder (Betz, Eickhoff, & Sullivan, 2013; Caesar & Kohler, 2009). In a recent consensus study using the Delphi technique, over 40 SLPs with expertise in the field of child language reached consensus on terminology for describing different language assessment practices (Denman, Kim, Munro, Speyer, & Cordier, 2019). These include terms relating to the type of data collected, with assessments described as being norm-referenced or criterion-referenced/descriptive; or terms relating to the type of tasks, with assessments described as de-contextualised, contextualised, or activity-focused. The environmental context in which abilities are assessed may be described as being either a clinical context, school context or home/community context. In addition, SLPs may incorporate dynamic procedures when assessing children, as opposed to using assessments that are purely static in nature. Dynamic procedures may be conducted as gradual prompting, test-teach-retest or both of these. See Table 5.1 for definitions of terms for describing the features of different types of language assessments.

Table 5.1

*Taxonomy terms (with definitions and examples) for describing assessments according to data types, task type, environmental context and dynamic features (from Denman et al., 2019)*

---

**Data Type (component of Taxonomy Aspect IV)**

---

**Norm-referenced**
Assessments that quantitatively compare a child's performance on a test or task to the performance of a sample of matched peers who completed the same test or task i.e. provide standard scores, percentile ranks or means and standard deviations. These assessments should always have standardised administration and scoring procedures.

---

**Criterion-referenced/descriptive**
Assessments that compare a child's performance against a pre-determined criterion such as developmental, grade or curriculum expectations (criterion-referenced); or are designed to give descriptive or qualitative data on a child's abilities (descriptive). These assessments may or may not have standardised administration and scoring procedures.

---

**Task Type (component of Taxonomy Aspect IV)**

---

**De-contextualised**
Discrete or 'pure' skills are assessed (which may be used to infer functional performance). Typically, the assessment occurs as a highly structured procedure directed by the assessor. Assessment items are administered as a series of questions that do not relate to a specific topic or situation. Examples:
- *Picture identification or naming tasks, sentence completion tasks, producing a sentence about a given picture, identifying or segmenting sounds in words, spelling individual words*

---

**Contextualised**
Abilities are assessed within a naturalistic communication context. The assessment is highly structured but occurs within a meaningful interaction, such as book reading or storytelling. If discrete skills are assessed, these are directly related to the communicative context being assessed. Examples:
- *Language sampling during conversation or play, narrative retelling, text comprehension tasks, role play tasks*
- *Parent or teacher interviews/checklists that document communication behaviours in specific communicative situations or contexts e.g. retelling events, asking for items, understanding jokes*

---

**Activity-Focused**
Abilities are assessed in relation to actual daily activities in which the child participates. The assessment is less structured with the assessor observing or being part of the interaction rather than directing tasks. If discrete skills are assessed, these are directly related to activity being completed. Examples:
- *Observing a child during free play with peers at lunchtime and noting abilities and behaviours*
- *Parent or teacher interviews/checklists that document the child's performance in, or level of support required to participate in, specific daily activities e.g. ordering food at a cafe, completing a class assignment, interacting with friends, following a recipe*

---

---

**Environmental Context (component of Taxonomy Aspect III)**

---

**Clinical context**

Abilities are examined within a clinical context i.e. the assessment does not incorporate materials or communication partners from the day-to-day environment. Examples:

- *Observing a child interacting in a clinical environment with clinical materials (e.g. materials the SLP has provided)*
- *Standardised tests conducted by the SLP using standardised test materials*

---

**School context**

Abilities are examined in a school (or kindergarten) context i.e. assessment incorporates communication partners, situations or materials that represent a school environment. Examples:

- *Interviewing a teacher regarding a child's communication at school*
- *Observing a child during regular class lesson and noting abilities and behaviours*
- *Assessing story retelling abilities using the actual book being studied in class*

---

**Home/other community context**

Abilities are examined in a home or community context i.e. assessment incorporates communication partners, situations or materials that represent a home/community environment. Examples:

- *Interviewing a parent regarding communication at home*
- *Observing a child interacting with a parent and siblings whilst playing with toys similar to those at home*

---

**Dynamic (component of Taxonomy Aspect IV)**

---

**Test-teach-retest**

Assessment designed to investigate a child's learning potential by evaluating response to explicit teaching. Examples:

- *Collecting a narrative retell, teaching the features of a good narrative story and then collecting another narrative retell in the same session*

---

**Gradual prompting**

Assessment designed to investigate a child's learning potential by evaluating the level of support the child requires to learn. Examples:

- *The child's ability to answer questions assessed using different question types and varied levels of prompting in order to determine the level of support the child needs to be successful*

---

Evidence-based practice recommendations stipulate that SLPs should use a range of procedures when assessing the language abilities of primary school children (Bishop, Snowling, Thompson, & Greenhalgh, 2016). Beyond this, there is limited consensus regarding the procedures that are most suited for assessing children with language disorder (Reilly et al., 2014). Policies outlining assessment practices exist at the level of individual states or service agencies, however these are typically associated with determining service

eligibility and may not be well-aligned with evidence-based practice recommendations that are outlined in literature (Spaulding, Szulga, & Figueroa, 2012). In previous studies conducted in the United States of America, SLPs in school settings have reported that guidelines for determining service eligibility vary considerably depending on state or local governments, with some SLPs reporting not having any such guidelines in their workplaces (Fulcher-Rood, Castilla-Earls, & Higginbotham, 2018; Hoffman, Ireland, Hall-Mills, & Flynn, 2013). As a consequence, language assessments practices may be chosen based on policy guidelines that are varied, clinical judgement or often, SLP preference (Fulcher-Rood et al., 2018; Roulstone, 2001; Singh, Chan, & Rusli, 2016).

As research evidence grows in the field of child language, there is increasing recognition of an urgent need to develop clinical practice guidelines (Fulcher-Rood et al., 2018; Roulstone, 2001). The field of implementation science recognises that numerous factors may influence the successful uptake of evidence-based practice recommendations into clinical practice, therefore it is important that consideration be given to the real-world contexts in which guidelines are to be implemented (Fulcher-Rood et al., 2018; Olswang & Prelock, 2015). As such, knowledge of the current status quo with regards to clinical child language assessment practice is needed. This includes information on the regularity with which SLPs currently use different types of assessments as well as an understanding of factors that influence use of different types of assessments, challenges experienced by SLPs in relation to child language assessment and main sources of information from which SLPs obtain information on assessment practice (Eadie, 2003).

Previous surveys have examined child language assessment practices however these surveys have predominantly focused on particular clinical populations, particular groups of SLPs or the use of specific types of assessment practices. For example, previous surveys have examined assessment for children from culturally and linguistically diverse (CALD) backgrounds or children with specific disabilities (Arias & Friberg, 2015; Caesar & Kohler,

2007, 2009; Teoh, Brebner, & McAllister, 2017; Watson & Pennington, 2015; Williams & McLeod, 2012). Other surveys have examined assessment practices of SLPs employed in schools in the United States of America (Beck, 1995; Caesar & Kohler, 2009; Fulcher-Rood et al., 2018; K. S. Wilson, Blackmon, Hall, & Elcholtz, 1991) or investigated SLPs use of a single type of assessment such as norm-referenced measures or contextualised procedures such as 'language sampling' (Betz et al., 2013; Huang, Hopkins, & Nippold, 1997; Kemp & Klee, 1997; Pavelko, Owens, Ireland, & Hahs-Vaughn, 2016; Westerveld & Claessen, 2014). These previous surveys provide valuable information on specific areas of child language assessment practice, however do not have the broad focus needed to understand the landscape of SLPs assessment practice. Further survey data are needed to examine the range of assessments that SLPs from different agencies use for a broad population of children (Fulcher-Rood et al., 2018).

Literature recognises that, across English-speaking countries, SLP services for primary school children are provided by a wide range of different service agencies with each agency having different jurisdictions, funding sources, policies and role descriptions (Rosenfeld, 2002; Spaulding et al., 2012; Speech Pathology Australia, 2014). Service availability and provision may also vary between metropolitan and regional/rural locations (Ruggero, McCabe, Ballard, & Munro, 2012). It is possible that the regularity with which different types of assessments are used may vary by service agency or geographical location, however this has not been investigated. To better understand SLP assessment practice, studies are needed to compare the language assessment practices used by SLPs from different agencies, different states or areas with differing remoteness classifications (Fulcher-Rood et al., 2018).

It is also possible that SLP assessment practice is influenced by SLP years of working experience or level of qualification. No previous studies have explicitly examined the influence of SLP level of qualification on assessment practice. Previous studies have

examined child language assessment practice in relation to SLPs years of working experience (Caesar & Kohler, 2007; Hux, Morris-Friehe, & Sanger, 1993; Pavelko et al., 2016; Roulstone et al., 2015), however findings are mixed. Although years of working experience has not been identified as a factor that directly influences the assessment practices SLPs use, it has been identified that differences in practice exist between SLPs with more or less years of experience. SLPs with more years of experience may be more likely to assess in different contexts, create their own assessment protocols or draw on their own judgement when interpreting assessment results (Caesar & Kohler, 2007; Pavelko et al., 2016; Roulstone et al., 2015; K. S. Wilson et al., 1991). This suggests that further surveys examining the broad range of language assessment practices that SLPs may use are needed to develop a deeper understanding of the influence of SLP level of qualifications and years of working experience on assessment practice.

As norm-referenced language measures are not suitable for children from CALD backgrounds (Arias & Friberg, 2015), it is expected that SLP assessment practices may vary depending on the proportion of children on SLP caseloads from CALD backgrounds. Only one previous study has examined SLP language assessment practices in relation to the proportion of children from CALD backgrounds on SLP caseload (Caesar & Kohler, 2007). Findings from this study did not identify differences between groups in relation to the types of assessments used by SLPs. However, as other studies have identified that assessment practice for children from CALD backgrounds may be changing over time (Arias & Friberg, 2015), there is a need for new studies examining current assessment practice.

Knowledge of the challenges SLPs experience in relation to child language assessment is important for understanding potential barriers and facilitators to implementation of evidence-based practice recommendations. Previous surveys have identified that limited time, budget constraints and limited access to training in conducting assessment may influence SLP language assessment practice (Arias & Friberg, 2015;

Fulcher-Rood et al., 2018; Westerveld & Claessen, 2014). Only one previous survey has examined the challenges reported by SLPs across different service agencies. The study, conducted by Huang et al. (1997) more than 20 years ago, only examined norm-referenced procedures. The authors identified that SLPs in education agencies were less satisfied with available norm-referenced measures and more likely to report challenges related to limited time and budget constraints compared with SLPs in other service agencies. Given that information from this survey is dated, further information is needed understand the current challenges experienced by SLPs across a varied range of work contexts.

Information on the sources from which SLPs obtain information on language assessment practices is also valuable in understanding the context that surrounds language assessment practice. Previous studies have reported that SLPs tend to rely on peers or workshops rather than journal articles for information on assessment practice (Beck, 1995; K. S. Wilson et al., 1991), however these studies were conducted over 20 years ago and may not be reflective of current practice. Obtaining information on the sources from which SLPs currently obtain information on assessment practice will assist in building an understanding of the context that surrounds language assessment practice and identify avenues for effective dissemination of future practice recommendations.

In summary, there is a need for survey data on the types of language assessments that SLPs use across a broad population of children. Investigations are needed to examine if factors such as service agency, geographical location, years of experience, SLP qualifications or proportion of children on SLP caseloads from CALD backgrounds influence SLP assessment practice. A better understanding is also needed regarding the challenges that SLPs experience in relation to language assessment and sources of information from which SLPs obtain information on assessment practices. This information will assist in profiling current practice and identifying future actions that may improve language assessment practice for primary school children (Roulstone, 2001).

**5.2.1. Objectives.** This study investigated the current assessment practices used by Australian SLPs when assessing the language abilities of primary school children aged 4-12 years. Specifically, the following objectives were addressed:

1. To identify the types of assessments SLPs use most regularly when assessing the language abilities of primary school children.

2. To identify if the following factors influence the regularity with which different types of language assessments are used by SLPs for primary school children: service agency, years of experience, SLP qualifications, proportion of children from CALD backgrounds on SLP caseload or geographical location in terms of Australian state and remoteness area classification.

3. To identify the challenges that SLPs in different agencies most frequently report in relation to child language assessment.

4. To identify the sources of information that SLPs from different agencies report most frequently obtaining information on child language assessment.

Descriptors from a taxonomy developed in the earlier consensus study by Denman et al. (2019) were used to describe types of language assessments. This taxonomy has four aspects across which language assessments may be described. In this current study, SLPs were asked to report on their assessment practices in relation to the environmental context assessed (component in Aspect III of the taxonomy), type of data collected in the assessment, type of tasks in the assessment and presence of dynamic features in the assessment (components of Aspect IV of the taxonomy). Survey data regarding the specific  language measures, assessment procedures and assessment delivery methods SLPs use to assess the language abilities of school-aged children has been reported on in an accompanying publication (Denman, Cordier, Munro, Kim, & Speyer, Under Review). The data in this accompanying publication relates to relates to Taxonomy Aspect I (assessment domains), Aspect II

(assessment purpose), the remaining component of Aspect III (assessment methods) and the remaining component of Aspect IV (presence of standardisation).

This study was restricted to one English-speaking country (Australia). This allowed for variables such as service agency and geographical location to be investigated in the absence of possible variations that may exist across countries. In most English-speaking countries services to school-aged children are provided by SLPs from different agencies and with differing levels of experience, qualifications and caseload characteristics (Rosenfeld, 2002; Roulstone, 2001; Speech Pathology Australia, 2014). Therefore, findings from this survey have relevance to SLPs internationally.

**5.3. Methods**

This study used an online survey created with Qualtrics software (Qualtrics, 2005) to collect data from SLPs regarding the types of language assessments they use. Ethical approval to conduct the survey was obtained from the Curtin University Human Research Ethics Committee (Approval Number: HRE2017-0659).

**5.3.1. Survey structure and format.** The survey was created with reference to literature on electronic survey design (Andrews, Nonnecke, & Preece, 2003; Oppenheimer, Pannucci, Kasten, & Haase, 2011). Questions were either 5-point Likert scale responses or multiple-choice questions with open text boxes for participants to add any response options that were not listed. The survey consisted of four sections. To assist in determining the size of the sample population, all Australian SLPs were eligible to complete the first section of the survey, regardless of their area of practice. This section of the survey consisted of questions about Speech Pathology Australia association membership, gender, postcode of workplace, years since graduation, qualifications and nature of current employment. The remaining survey sections were completed by SLPs who indicated that they provided a service to at least 40 children with language disorder in the last year. In the second section, SLPs were asked questions about the service agency in which they work and the proportion of children

on their caseload from CALD backgrounds. SLPs were also asked to indicate the main (up to

four) challenges they experience in relation to assessment for school-aged children with

language disorder and the main (up to three) sources from which they obtain information

about child language assessment practices. The third section consisted of Likert scale

questions regarding the frequency with which SLPs used different types of assessments. A

copy of the survey questions relating to this study (sections one to three) is supplied in

Supplementary Appendix 5.1. The questions in fourth section of the survey related to the

names of actual language measures used and results from this section are reported in the

accompanying publication.

To ensure consistent descriptions of assessment practices between participants,

careful consideration was given to the terminology used within the survey questions. The

terms used were those agreed upon in a previous consensus study (Denman et al., 2019) and

participants were instructed to apply these definitions when answering questions, even if they

use the terms differently themselves. See Table 5.1 for a list of the terms and definitions used

in the survey. All terms and definitions in the survey were accompanied by examples of

assessments that are described by each term. A supplementary table with a list of assessments

already categorised using taxonomy terminology was also provided for participants to refer to

during the survey if needed. This supplement is provided in Supplementary Appendix 5.2.

To ensure consistent application of the frequency rating scale by survey participants,

the Likert scale points were associated with descriptors, as well as numeric qualifiers relating

to the proportion of children assessed using each procedure (Blais & Grondin, 2011). For

example, participants were asked "How many children were assessed in a school context

(considering the last 40 children assessed)?" rated on a Likert scale of most = 34 or more

children, many = between 20-34 children, some = between 6-19 children and few = less than

5 children and none = no children. The reference number of 40 was selected because it was

considered large enough to capture trends, but still small enough for participants to recall the types of assessments they used.

Prior to the survey being distributed, four SLPs piloted the survey and provided feedback. These participants were sourced through email listservs and the professional networks of the researchers. Pilot participants were from the Australian states of Queensland or New South Wales and were all from different service agencies, including a public education (school) service, private practice, a non-government disability service agency and a university clinic.

**5.3.2. Survey dissemination**. The survey was open for four months between mid-February and mid-June 2018 and was advertised through the Speech Pathology Australia national newsletter sent monthly to all association members. The survey link was also circulated on Twitter, Facebook and emailed through professional networks of the researchers, publicly available email addresses and email discussion groups. SLPs who received the survey link were asked to forward the link to colleagues. The survey was estimated to take 5 minutes for SLPs who only completed the first section and between 25-40 minutes for SLPs who completed all four sections. Participants were able to complete the survey in more than one sitting as the survey could be saved and opened up again at a later time.

**5.3.3. Data analysis**. Survey responses were imported into the Statistical Package for the Social Sciences (SPSS version 20 program (IBM Corp, Released 2011). State and remoteness area classification were assigned from the postcodes provided by survey participants. Remoteness area was classified by the Australian Statistical Geography Standard (ASGS) developed by the Australian Bureau of Statistics (Australian Bureau of Statistics, 2016). Remoteness area was collapsed into two categories 'major city' (ACGS category of major city) and 'regional-remote' (AGCS categories of inner regional, outer regional, remote and very remote).

Descriptive statistics were used to examine the frequency with which SLPs use each type of assessment, the main challenges reported by SLPs, and the main sources of information on assessment reported by SLPs. For multiple choice options, responses supplied by participants in open text boxes were coded to an existing response option if applicable or coded as a new response option. Chi Square tests were used investigate differences between groups of SLPs. Backward elimination binary logistic regression analyses were used to investigate the factors that influence the frequency with which each type of assessment was used by SLPs (Sperandei, 2014).

To create binary dependent variables for regression analysis, Likert scale responses were transformed into variables with two response categories: regularly (options 'many' or 'most') or not regularly (options 'none', 'few', 'and 'some'). This means that regular use of an assessment was identified if an SLP reported using the procedure with 20 or more of the last 40 children assessed (i.e. half or more children). Independent variables were state, service agency, years since graduation, SLP qualifications, remoteness area classification and proportion of children from CALD backgrounds. Due to the small numbers of survey participants in the states of Tasmania, Northern Territory and Australian Capital Territory, the 32 participants from these three states and territories were removed from regression analysis to improve the sensitivity (Sperandei, 2014). Consequently, the sample size was n=407 for all analyses except regression analysis where sample size was n= 375.

To reduce the initial number of variables in the multivariate regression analyses, a pre-selection process was employed by using a series of univariate Chi-square tests (Sperandei, 2014). Only variables with $X^2$ p-value of less than 0.1 in the univariate pre-selection were included in the multivariate logistic regression models. The variables that best contributed to the regression models were then identified through a backward elimination process. This occurred by conducting logistic regression with all the pre-selected variables

and removing non-significant variables one by one (starting with the least significant variable) until only the variables that significantly contributed to the model remained.

Reference groups are the categories in each variable to which other categories are compared for statistical significance during regression analysis (Sperandei, 2014). The reference groups in this analysis were 'New South Wales (NSW)' for state, 'private practice' for service agency, '0-2 years' for years since graduation, 'Bachelor of SLP with no additional qualifications' for SLP qualifications and 'major city' for remoteness area classification. Reference groups were chosen as groups with the largest sample size, or in the case of years since graduation, for ease of interpretation by taking the lowest category in the scale (Sperandei, 2014). NSW was selected as the reference group as this state had both a large sample size and the most evenly distributed population across subcategories of other variables, particularly service agency.

## 5.4. Results

**5.4.1. Survey responses.** In total, 847 SLPs consented to take part in the survey, with 727 providing complete and valid survey responses (85.8% completion rate). Of the SLPs who completed the survey, 83.4% identified themselves as being members of the national speech pathology association (Speech Pathology Australia). This figure is comparable with the 80% estimate obtained in a previous survey of Australian SLPs (Westerveld & Claessen, 2014). Personal communication with Speech Pathology Australia about its membership database indicated that approximately 53% of qualified Australian SLPs who are members of the association, work with primary school children (L. Young, personal communication, 4th June and 20th September, 2018). Using 83% as an estimate of association membership and 53% as an estimate of the proportion of SLPs who work with children aged 4-12; we calculated that 4,610 SLPs in Australia work with children aged 4-12 years. In this survey, 525 SLPs identified themselves as working with children aged 4-12 years, with this response rate representing 11.4% of the estimated population. As data on association membership were

also available per state, it was approximated that between 7.7% and 40.1% of the estimated number of SLPs in each Australian state who work with primary school children completed the survey. The number of SLPs surveyed in relation to the estimated population for each state/territory is provided in Supplementary Appendix 5.3.

An estimate of the number of Australian SLPs who frequently work specifically with primary school children with language disorder was not available from the membership database. Of the 525 participants who worked with primary school aged children, 407 (77.5%) identified themselves as regularly providing clinical services to this population, as defined by having provided a service to 40 or more primary school aged children with language disorder in the last 12 months. These 407 participants were the sample of interest in this survey.

**5.4.2. Participant demographics**. Participant characteristics are outlined in Table 5.2. Approximately two thirds of participants (63.1%) reported working full-time and one third (36.8%) reported working part-time. Approximately one third of SLPs (33.2%) worked in education agencies (i.e., government or non-government school or education service) and another third (31.1%) worked in private practice (i.e., business owner or employee in private practice). The remaining proportion worked in health agencies (i.e., government or non-government hospital or health service) (15.7%), disability-specific agencies (i.e., government or non-government agency with eligibility criteria that children must have a disability or suspected disability; 14.0%), general agencies (i.e., government or non-government agency that is not education, health or disability-specific) and university (i.e., student teaching clinics; 2.2%). There was a wide spread amongst participants with regards to years since graduation and qualifications. One third (30.5%) of participants worked in agencies based outside major cities and 15.0% reported that more than half of their caseload comprised of children from CALD backgrounds.

Table 5.2

*Demographics of Survey Participants Who Work with Children Aged 4-12 Years with Language Disorder (Survey Part I; n=407)*

| Category | Subcategory | Australian State or Territory of Workplace | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Australian Capital Territory (n=11) | New South Wales (n=103) | Northern Territory (n=7) | Queensland (n=116) | South Australia (n=36) | Tasmania (n=14) | Victoria (n=71) | Western Australia (n=49) | Total (n=407) |
| Gender | Female | 10 (90.95%) | 101 (99.5%) | 7 (100%) | 111 (95.7%) | 35 (97.2%) | 14 (100%) | 69 (97.2%) | 46 (93.9%) | 393 (96.6%) |
| | Male | 1 (9.1%) | 2 (3.3%) | 0 (0.0%) | 4 (3.4%) | 1 (2.8%) | 0 (0.0%) | 2 (2.8%) | 3 (6.1%) | 13 (3.2%) |
| | Other | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1 (0.3%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1 (0.2%) |
| Employment Status | Work full-time (1 FTE) | 9 (81.8%) | 69 (67.0%) | 6 (85.7%) | 73 (62.9%) | 23 (63.9%) | 8 (57.1%) | 44 (62.0%) | 25 (51.0%) | 257 (63.1%) |
| | Work part-time (< 1 FTE) | 2 (18.2%) | 34 (33.0%) | 1 (14.2%) | 43 (37.1%) | 13 (36.1%) | 6 (42.8%) | 27 (38.0%) | 24 (48.9%) | 150 (36.8%) |
| Agency through which SLP service is provided | [a] Education agency | 1 (9.1%) | 5 (4.9%) | 1 (14.3%) | 64 (55.2%) | 22 (61.1%) | 12 (85.7%) | 23 (32.4%) | 7 (14.2%) | 135 (33.2%) |
| | [b] Private practice | 7 (63.6%) | 39 (37.9%) | 1 (14.3%) | 5 (4.3%) | 6 (16.7%) | 2 (14.2%) | 21 (29.6%) | 27 (55.1%) | 129 (31.1%) |
| | [c] Health agency | 1 (9.1%) | 35 (34.0%) | 2 (28.6%) | 5 (4.3%) | 1 (27.7%) | 0 (0.0%) | 13 (18.3%) | 7 (14.2%) | 64 (15.7%) |
| | [d] Disability specific | 2 (18.2%) | 14 (13.6%) | 0 (0.0%) | 11 (9.2%) | 5 (13.9%) | 0 (0.0%) | 11 (15.5%) | 5 (10.2%) | 48 (14.0%) |
| | [e] General agency | 0 (0.0%) | 5 (4.9%) | 2 (18.2%) | 9 (7.8%) | 1 (27.7%) | 0 (0.0%) | 2 (2.8%) | 3 (7.5%) | 22 (5.4%) |

**Australian State or Territory of Workplace**

| Category | Subcategory | Australian Capital Territory (n=11) | New South Wales (n=103) | Northern Territory (n=7) | Queensland (n=116) | South Australia (n=36) | Tasmania (n=14) | Victoria (n=71) | Western Australia (n=49) | Total (n=407) |
|---|---|---|---|---|---|---|---|---|---|---|
| Years since graduation | 0-2 years | 5 (45.5%) | 16 (15.5%) | 0 (0.0%) | 21 (18.1%) | 4 (11.1%) | 1 (7.1%) | 8 (11.3%) | 5 (10.2%) | 60 (14.7%) |
| | 3-5 years | 5 (45.5%) | 25 (24.2%) | 1 (14.2%) | 35 (30.2%) | 11 (30.5%) | 6 (42.9%) | 12 (16.9%) | 9 (18.4%) | 104 (25.6%) |
| | 6-10 years | 1 (9.1%) | 22 (21.4%) | 1 (14.2%) | 16 (13.8%) | 7 (19.4%) | 3 (21.4%) | 17 (23.9%) | 9 (18.4%) | 76 (18.7%) |
| | 11-20 years | 0 (0.0%) | 21 (20.4%) | 2 (28.6%) | 23 (19.8%) | 4 (11.1%) | 3 (21.4%) | 22 (31.0%) | 12 (24.4%) | 87 (21.4%) |
| | 21+ years | 0 (0.0%) | 19 (18.4%) | 3 (42.9%) | 21 (18.1%) | 10 (27.8%) | 1 (7.1%) | 12 (16.9%) | 14 (28.6%) | 80 (19.7%) |
| Qualifications | Other qualification/s besides Bachelor's degree | 6 (54.5%) | 38 (39.1%) | 4 (57.1%) | 38 (32.8%) | 10 (27.8%) | 7 (50.0%) | 26 (36.6%) | 16 (32.7%) | 145 (35.6%) |
| | Bachelor's degree with honours | 1 (9.1%) | 6 (6.2%) | 0 (0.0%) | 14 (12.1%) | 2 (55.5%) | 1 (7.1%) | 5 (7.0%) | 6 (12.2%) | 35 (8.6%) |
| | Bachelor's degree | 4 (36.3%) | 59 (60.8%) | 3 (42.8%) | 64 (55.2%) | 24 (66.7%) | 6 (42.8%) | 40 (56%) | 27 (55.1%) | 227 (55.8%) |
| [g]Remoteness (ACGS classification) | Major City | 11 (100%) | 72 (69.9%) | 0 (0.0%) | 73 (62.9%) | 26 (72.2%) | 0 (0.0%) | 54 (67.1%) | 47 (95.9%) | 283 (69.5%) |
| | Regional-Remote | 0 (0.0%) | 31 (30.1%) | 7 (100%) | 43 (37.1%) | 10 (27.8%) | 14 (100%) | 17 (23.9%) | 2 (4.1%) | 124 (30.5%) |
| [h]Proportion of children from CALD Backgrounds | Less than half of last 40 children | 11 (100%) | 76 (73.8%) | 5 (71.4%) | 106 (91.4%) | 32 (88.8%) | 14 (100%) | 55 (77.5%) | 47 (95.1%) | 346 (85.0%) |
| | More than half of last 40 children | 0 (0.0%) | 27 (26.2%) | 2 (28.6%) | 10 (8.6%) | 4 (11.1%) | 0 (0.0%) | 16 (22.5%) | 2 (4.1%) | 61 (15.0%) |

Note: [a] Education service or school (may be government or non-government); [b] Private practice i.e. business owner or employee in private practice; [c] Health service or hospital (may be government or non-government); [d] Disability agency i.e. children must have diagnosis (or suspected diagnosis) of disability to access the service (may be government or non-government); [e] Agency that is not identified as education, health or disability specific (may be government or non-government); [f] University clinic i.e. student teaching clinic; [g] As classified by Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS)*, 2016 [cited 2018 March] available from:

http://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.005.

## 5.4.3. Regularity of assessment use and factors that influence regularity of use.

The percentage of SLPs who reported regularly using each type of assessment is displayed in Figure 5.1.
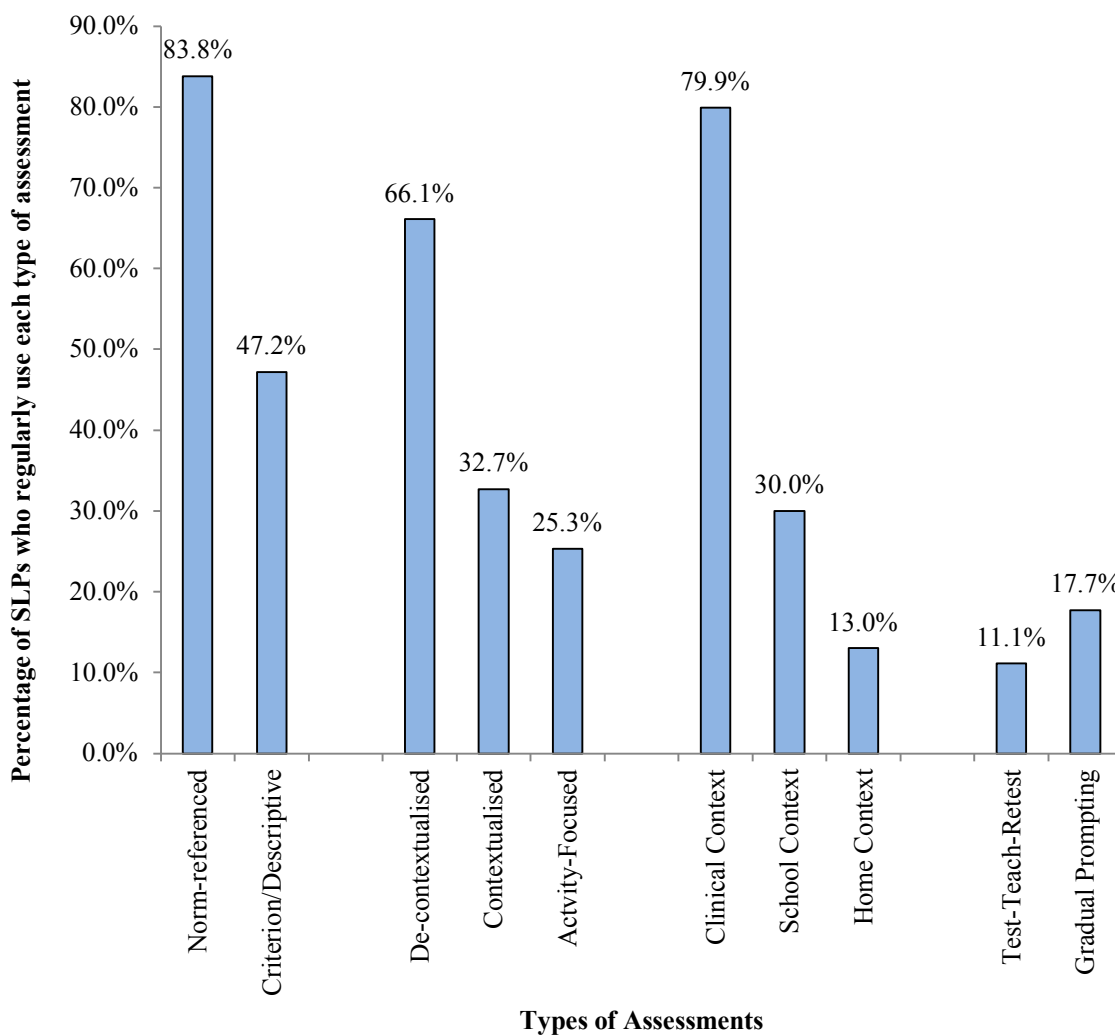


*Figure 5.1.* Percentage of SLPs who reported regularly using each type of assessment (n=407)
Regular use was defined as being used with half or more of the last 40 children who were assessed. Assessments are described by: Data Type (each assessment is either norm-referenced or criterion-referenced/descriptive), Task Type (each assessment is either de-contextualised, contextualised or activity-focused), Environmental Context (each assessment targets either a clinical, school or home/other community context) and Dynamic (assessments may or may not have a dynamic component: Test-Teach-Retest or Gradual Prompting).

Results of univariate variable pre-selection are displayed in Table 5.3. Variables found to be significant in univariate pre-selection were selected for inclusion in the subsequent multivariate regression analysis. Results of multivariate regression analysis are displayed in Table 5.4.

Table 5.3

*Univariate Analysis: Variables that Influence the Regularity with which Different Types of Assessments are Used (n=375)*

| Dependent Variable | Independent Variable (factors that influence assessment use) | | | | | |
|---|---|---|---|---|---|---|
| Assessment type | Agency | State | Years since graduation | SLP Qualifications | [a] Remoteness (ACGS classification) | [b] Number of children from CALD backgrounds |
| Norm-referenced (Data Type) | <0.001** | 0.084* | 0.162 | 0.721 | 0.183 | 0.356 |
| Criterion referenced (Data Type) | 0.454 | 0.067* | 0.211 | 0.450 | 0.503 | 0.393 |
| De-contextualised (Task Type) | <0.001** | 0.955 | 0.432 | 0.109 | 0.683 | 0.920 |
| Contextualised (Task Type) | 0.019** | <0.001** | 0.007* | 0.312 | 0.004** | 0.303 |
| Activity-focused (Task Type) | <0.001** | 0.056* | 0.496 | 0.624 | 0.711 | 0.503 |
| Clinical Context (Environmental Context) | <0.001** | 0.539 | 0.445 | 0.948 | 0.373 | 0.595 |
| School Context (Environmental Context) | < 0.001** | 0.025* | 0.689 | 0.918 | 0.917 | 0.780 |
| Home/Other Context (Environmental Context) | 0.047* | 0.018* | 0.433 | 0.157 | 0.061* | 0.903 |
| Test-Teach-Retest (Dynamic) | 0.334 | <0.001** | 0.228 | 0.039* | 0.166 | 0.917 |
| Gradual prompting (Dynamic) | 0.264 | 0.003** | 0.008* | 0.355 | 0.238 | 0.129 |

Note: *$p$<0.1 (variables with p<0.1were selected for inclusion in multivariate regression analyses); ** $p$<0.005; [a] As classified by Australian Bureau of Statistics: *Australian Statistical Geography Standard (ASGS)*, 2016 [cited 2018 March] available from: http://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.005; [b] Number of children (considering the last 40 children assessed) identified as having Cultural and Linguistic Diversity (CALD) e.g. bilingualism or standard Australian English is not child's first language.

Table 5.4

*Multivariate Regression Models: Factors that Influence the Regularity with which Different*

*Types of Assessments are Used (n=375)*

**Norm-referenced: Nagelkerke $R^2$= 12.1; p-value <0.001**

| Independent variable (n) | % of SLPs[a] | p-value | Odds Ratio | 95% CI for odds ratio |
|---|---|---|---|---|
| Agency | | <0.001*** | | |
| Private practice (118) | 88.1 | Reference | 1 (reference) | |
| Disability (46) | 65.2 | 0.001** | 0.252[b] | 0.111 – 0.576 |
| Education (121) | 89.3 | 0.784 | | |
| General (20) | 50.0 | <0.001*** | 0.135[b] | 0.048 – 0.380 |
| Health (58) | 88.5 | 0.939 | | |
| University (9) | 88.9 | 0.946 | | |

**De-contextualised: Nagelkerke $R^2$= 9.5; p-value <0.001**

| Independent variable (n) | % of SLPs[a] | p-value | Odds Ratio | 95% CI for odds ratio |
|---|---|---|---|---|
| Agency | | <0.001*** | | |
| Private practice (118) | 72.9 | Reference | 1 (reference) | |
| Disability (46) | 45.7 | 0.001** | 0.31[b] | 0.15 - 0.63 |
| Education (121) | 75.2 | 0.682 | | |
| General (20) | 30.0 | 0.001** | 0.16[b] | 0.06 - 0.45 |
| Health (58) | 67.2 | 0.429 | | |
| University (9) | 55.6 | 0.276 | | |

**Contextualised: Nagelkerke R2= 15.2; p-value <0.001**

| Independent variable (n) | % of SLPs[a] | p-value | Odds Ratio | 95% CI for odds ratio |
|---|---|---|---|---|
| State | | <0.001*** | | |
| New South Wales (103) | 26.2 | reference | 1 (reference) | |
| Queensland (116) | 42.2 | 0.007** | 2.24 | 1.24 - 4.04 |
| South Australia (36) | 25.0 | 0.818 | | |
| Victoria (71) | 21.1 | 0.327 | | |
| Western Australia (49) | 61.2 | <0.001*** | 4.21 | 2.01-8.83 |
| | | | | |
| Years Since Grad | | 0.008** | | |
| 0-2 years (54) | 16.7 | Reference | 1 (reference) | |
| 3-5 years (92) | 30.4 | 0.061 | | |
| 6-10 years (71) | 35.2 | 0.010* | 3.26 | 1.32 - 8.00 |
| 11-20 years (82) | 40.2 | 0.003** | 3.81 | 1.59 - 9.14 |
| 21+ years (76) | 46.1 | 0.001** | 4.55 | 1.89 - 10.96 |

**Activity-Focused: Nagelkerke R2= 25.1; p-value <0.001**

| Independent variable (n) | % of SLPs[a] | p-value | Odds Ratio | 95% CI for odds ratio |
|---|---|---|---|---|
| Agency | | <0.001*** | | |
| Private practice (118) | 4.9 | Reference | 1 (reference) | |
| Disability (46) | 63.0 | <0.001*** | 15.07 | 6.47 - 35.20 |
| Education (121) | 33.9 | 0.239 | 4.53 | 2.24 - 9.17 |
| General (20) | 35.0 | <0.001*** | 4.76 | 1.59 - 14.23 |
| Health (58) | 10.2 | 0.005** | | |
| University (9) | 33.3 | 0.054 | 4.42 | 0.98-19.97 |

**Clinical Context: Nagelkerke $R^2$= 14.3; p-value <0.001**

| Independent variable (n) | % of SLPs[a] | p-value | Odds Ratio | 95% CI for odds ratio |
|---|---|---|---|---|
| Agency | | <0.001*** | | |
| Private practice (118) | 85.6 | reference | 1 (reference) | |
| Disability (46) | 54.3 | <0.001*** | 0.20[b] | 0.09 - 0.44 |
| Education (121) | 85.1 | 0.918 | | |
| General (20) | 50.0 | 0.001*** | 0.17[b] | 0.06 - 0.47 |
| Health (58) | 91.8 | 0.236 | | |
| University (9) | 77.8 | 0.530 | | |

**School Context: Nagelkerke $R^2$= 24.8; p-value <0.001**

| Independent variable (n) | % of SLPs[a] | p-value | Odds Ratio | 95% CI for odds ratio |
|---|---|---|---|---|
| Agency | | <0.001*** | | |
| Private practice (118) | 19.5 | reference | 1 (reference) | |
| Disability (46) | 43.5 | 0.002*** | 3.18 | 1.52 - 6.66 |
| Education (121) | 53.7 | <0.001** | 4.79 | 2.69 - 8.55 |
| General (20) | 5.0 | 0.147 | | |
| Health (58) | 4.9 | 0.015* | 0.21[b] | 0.06 - 0.74 |
| University (9) | 33.3 | 0.330 | | |

**Dynamic: Test-Teach-Retest: Nagelkerke $R^2$= 13.2; p-value <0.001**

| Independent variable (n) | % of SLPs[a] | p-value | Odds Ratio | 95% CI for odds ratio |
|---|---|---|---|---|
| State | | <0.001*** | | |
| New South Wales (103) | 14.6 | reference | 1 (reference) | |
| Queensland (116) | 10.3 | 0.345 | | |
| South Australia (36) | 2.8 | 0.09 | | |
| Victoria (71) | 1.4 | 0.018* | 0.08[b] | 0.01 - 0.65 |
| Western Australia (49) | 28.6 | 0.043* | 2.35 | 1.03 - 5.37 |

Note: Models were not significant (*p*<0.01) for criterion-referenced/descriptive or home/community context assessment; *p<0.05; **p<0.01; ***p<0.001; [a] Percentage of SLPs from each category who reported regularly

using each type of assessment; [b] In the text of the publication, categories with odds ratios less than 1.0 are reported as the corresponding ratio above 1.0 (inverse of the odds ratio).

*5.4.3.1. Data Type (norm-referenced or criterion-referenced/descriptive).* The majority of SLPs (83.8%) reported regularly using norm-referenced language measures (regular use being defined as being used with half or more of last 40 children). Only five participants (1.2%) indicated that they had not used norm-referenced language measures for any of the last 40 children they assessed; all these SLPs were from disability agencies. In comparison, 47.2% of SLPs indicated that they regularly use assessments that yield criterion-referenced/descriptive data (regular use being defined as being used with half or more of last 40 children). Eight participants (2.0%) indicated not having used criterion-referenced/descriptive assessments for any of the last 40 children they assessed, with these participants representing a variety of agencies.

Results of multivariate regression analysis indicated that the frequency with which SLPs use norm-referenced measures was influenced by service agency, with this factor accounting for 12.1% of the variance. Fewer SLPs in general agencies (50.0% of SLPs) and disability agencies (65.2% of SLPs) reported regularly using norm-referenced measures compared to with 88.1% or more of SLPs in other agencies. When these findings are reported in terms of odds ratios, SLPs in private practice had 7.41 times greater odds than SLPs in general agencies of reporting regular use of norm-referenced measures and 3.96 times greater odds than SLPs in disability agencies of reporting regular use of norm-referenced measures. No factors were identified as significantly influencing the regularity with which criterion-referenced/descriptive assessments were used.

*5.4.3.2. Task Type (de-contextualised, contextualised and activity-focused).* Two thirds (66.1%) of SLPs reported regular use of de-contextualised assessments, one third (32.7%) indicated regular use of contextualised assessments and only one quarter (25.3%) indicated regularly using activity-focused assessments (regular use being defined as being

used with half or more of last 40 children). Four participants (1.0%) reported not having used any de-contextualised assessments, 28 (6.9%) reported not having used contextualised assessments and 62 (15.2%) reported not having used any activity-focused assessments (considering the last 40 children they assessed).

Results of multivariate regression analysis indicated that the regularity with which de-contextualised assessments were used was influenced by service agency, with this factor explaining 9.5% of the variance. SLPs in general agencies (30.0% of SLPs) and disability agencies (45.7% of SLPs) were less likely to report regular use of these assessments, whilst SLPs in education (75.9% of SLPs) and private practice (72.9% of SLPs) were more likely. SLPs in private practice had 6.29 times greater odds than SLPs in general of reporting regular use of de-contextualised assessments and 3.19 times greater odds than SLPs in disability agencies of reporting regular use of de-contextualised assessments.

Use of contextualised assessments was influenced by state and years since graduation, with these two variables explaining 15.2% of the variance. The percentage of SLPs in Western Australia who reported regular use of contextualised assessments was 61.2% compared to 42.2% in Queensland and 26.2% or less in other states. In terms of odds ratios, SLPs in Western Australia had 4.21 times greater odds than SLPs in New South Wales of reporting regular use of contextualised assessments and SLPs in Queensland had 2.24 times greater odds than SLPs in New South Wales of reporting regular use of contextualised assessments.

Use of contextualised assessments also increased with increasing number of years since graduation. The percentage of SLPs with more than 20 years since graduation who reported regular use of contextualised assessment was 46.1% compared with 35.2% of SLPs with 6-10 years of experience and 16.7% of SLPs with two years or less experience. The odds ratios indicate that SLPs with more than 21 years since graduation had 4.55 times greater odds than SLPs with less than two years since graduation of reporting regular use of

contextualised assessments. SLPs with 6-10 years since graduation had 2.24 times greater odds than SLPs with two years or less since graduation of reporting regular use of contextualised assessments.

With regards to activity-focused assessments, service agency explained 25.1% of the variance in regularity of use. The percentage of SLPs in disability agencies reporting regular use of activity-focused assessments was 63.0% compared to 33.3-35.0% of SLPs in universities, education agencies and general agencies, 10.2% of SLPs in private practice and 4.9% in health agencies. SLPs in disability agencies had 15.07 times greater odds than of SLPs in private practice of reporting frequent use of activity-focused assessments, while SLPs in universities, education agencies and general agencies had approximately four times greater odds than SLPs in private practice of reporting frequent use of activity-focused assessments.

*5.4.3.3. Environmental context (clinical, school or home/community contexts).*
Most SLPs (79.9%) indicated regularly conducting assessment in a clinical context; 30.0% indicated regularly conducting assessment in a school context and only 13.0% indicated regularly conducting assessment in a home/community context (regular use being defined as being used with half or more of last 40 children). The number of SLPs who reported not conducting assessment in a clinical context was 20 (4.9%), school context was 76 (18.7%) and home/community context was 174 (42.8%).

Results of multivariate regression analysis indicated that the regularity with which assessments were conducted in a clinical context was influenced by service agency, with this factor accounting for 14.3% of the variance. Half of SLPs in general agencies (50.0%) and disability agencies (54.3%) reported regularly conducting assessment in a clinical context compared with 77.8% or more of SLPs in other agencies. SLPs in private practice had 5.95 times greater odds than SLPs in general agencies of reporting regular use of clinical context

assessments and 5.00 times greater odds than SLPs in disability agencies of reporting regular use of clinical context assessments.

The regularity with which assessments were conducted in a school context was also influenced by service agency, with this factor explaining 24.8% of the variance. The percentage of SLPs in education and disability agencies who reported regularly conducting assessments in a school context was 53.7% and 43.5% respectively, compared with 5.0% or less of SLPs in health or general agencies. The odds of SLPs in education agencies reporting regularly conducting assessment in a school context was 4.79 times greater than the odds of SLPs in private practice. The odds of SLPs in disability agencies reporting regularly conducting assessment in a school context was 3.18 times greater than the odds of SLPs in private practice. SLPs in private practice had 4.67 times greater odds than SLPs in health agencies of reporting regular use of school context assessments. No factors were identified as significantly influencing the regularity with which home/community context assessments were used.

*5.4.3.4. Dynamic (Test-teach-retest or gradual prompting).* Only 11.1% of SLPs reported regularly using dynamic-test-teach-retest assessments and only 17.7% reported regularly using dynamic-gradual prompting assessments (regular use being defined as being used with half or more of last 40 children). A total of 169 (41.5%) SLPs reported not using any dynamic-test-teach-retest assessments and 106 (26.0%) SLPs indicated not having used any dynamic-gradual prompting assessments.

Results of multivariate regression analysis indicate that use of dynamic test-teach-retest procedures were found to be influenced by state, with this variable explaining 13.2% of the variance. SLPs in Western Australia were more likely to regularly use this assessment (28.6% of SLPs) compared to 14.6% in New South Wales, 10.3% in Queensland and 2.8% or less in Victoria and South Australia. The odds of SLPs in Western Australia reporting regular use of dynamic test-teach-retest assessment was 2.35 times greater than the odds of SLPs in

New South Wales. The odds of SLPs in New South Wales reporting regular use of dynamic test-teach-retest assessment was 11.90 times greater than the odds of SLPs in Victoria.

Dynamic gradual-prompting assessments were influenced by both state and years since graduation, with these two variables accounting for 12.3% of the variance. SLPs in Western Australia were more likely to report regular use of these assessments (38.8% of SLPs) compared with 16.5% or less in other states. The odds of SLPs in Western Australia reporting regular use of dynamic gradual prompting assessment was 3.28 times greater than the odds of SLPs in New South Wales. SLPs with more than two years since graduation were also more likely to report regular use of dynamic gradual prompting assessments compared to SLPs with two years or less since graduation. The percentage of SLPs with more than two years since graduation who reported regular use of this procedure was 11.3% or more compared with 5.6% of SLPs with two years or less since graduation. The group with the highest percentage of SLPs reporting regular use of dynamic gradual prompting assessments were SLPs with 3-5 years since graduation (27.2% of SLPs). The odds of these SLPs reporting regular use of dynamic gradual prompting assessments was 6.71 times greater than the odds of SLPs with two years or less since graduation.

**5.4.4. Challenges reported by SLPs**. The most frequently reported challenges related to child language assessment related to lack of time: 'limited time to plan or analyse assessment' (35.1% of SLPs), 'limited time to meet with teachers' (33.4% of SLPs), 'limited time to meet with parents' (26.5% of SLPs), and 'limited face-to-face time with children for assessment' (22.1% of SLPs). Other frequently reported challenges included 'limited assessment materials' (e.g., due to budget constraints) (35.4% of SLPs), 'setting constraints' (i.e., not able to see children in particular locations) (15.5% of SLPs), 'workplace requirements' (i.e., workplace requires particular data or use of particular tools) (12.0% of SLPs). SLPs also reported challenges related to professional development needs such as 'lack of SLP skills or confidence with assessing children from CALD backgrounds' (23.6% of

SLPs), 'lack of skills or confidence with assessing complex needs' (17.9% of SLPs), 'lack of skills or confidence assessing literacy' (i.e., reading and writing; 15.0% of SLPs) and 'limited access to training in assessment' (16.7% of SLPs). The percentage of SLPs who identified each of these challenge as a main challenge is also displayed in Figure 5.2.
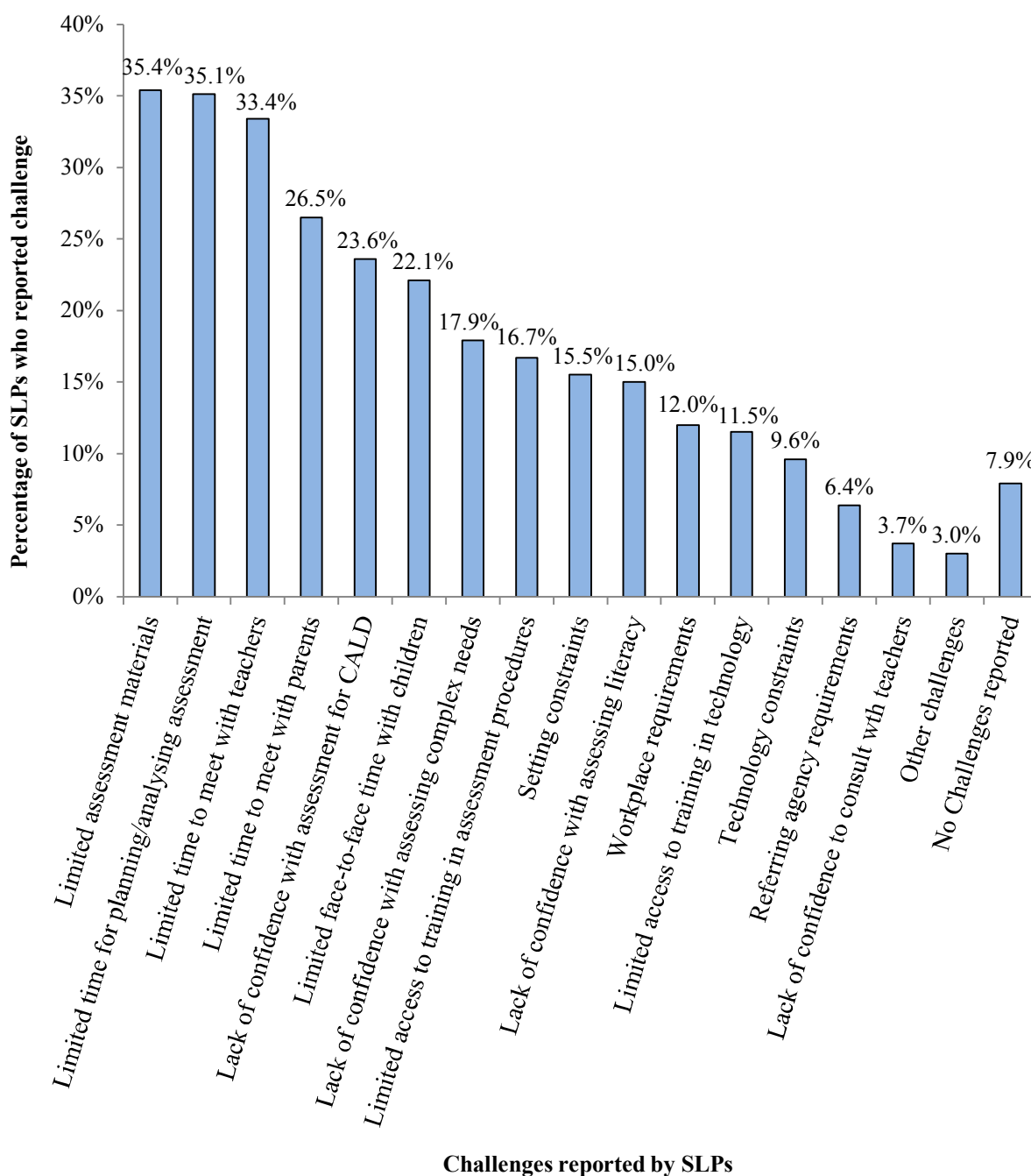


*Figure 5.2.* Percentage of SLPs who identified each challenge in relation to child language assessment (n=407). SLPs were able to select up to four main challenges.

Comparisons between groups indicated significant differences between SLPs from different agencies with regards to challenges reported. 'Limited time for planning and analysing assessment' was more likely to be reported by SLPs in universities (55.6%) and disability agencies (43.8%), compared with SLPs in private practices (21.7%) and health agencies (18.2%) $X^2$ (5, N=407) = 11.72, p =0.039. 'Limited time to meet with teachers' was more likely to be reported by SLPs in education agencies (43.8%), compared with SLPs in general (21.7%) and disability agencies (14.6%) $X^2$ (5, N=407) = 16.60, p =0.005. 'Limited time to meet with parents' was more likely to be reported by SLPs in universities (55.6%) and education agencies (45.2%), compared with SLPs in health agencies (12.5%) and private practice (14.7%) $X^2$ (5, N=407) = 44.78, p <0.001. 'Setting constraints' (i.e., not able to see children in particular locations) was more likely to be reported by SLPs in health agencies (31.2%) and universities (22.2%), compared with SLPs in general agencies (13.6%), disability agencies (12.5%) or education agencies (8.9%) $X^2$ (5, N=407) = 17.35, p =0.004. 'Workplace requirements' (i.e., workplace requires particular data or use of particular tools) were more likely to be reported by SLPs in disability (22.9%) and education agencies (21.5%), compared with SLPs in universities (0.0%) and general agencies (0.0%) $X^2$ (5, N=407) = 31.10, p <0.001. 'Limited assessment materials' (e.g., due to budget constraints) were more likely to be reported as a main challenge by SLPs in general agencies (54.5%) and disability agencies (50.0%), compared with SLPs in universities (22.2%) and education agencies (21.4%) $X^2$ (5, N=407) = 29.63, p <0.001.

SLPs who graduated more recently were more likely to report challenges related to 'lack of skills or confidence with assessing complex needs' and 'lack of skills or confidence with assessing literacy' (i.e., reading and writing). The percentage of SLPs with two or less years since graduation who reported 'lack of skills or confidence with assessing complex needs' was significantly higher (35.0%) than of SLPs with 6-10 years since graduation (17.1%) and of SLPs with 21 or more years since graduation (8.2%) $X^2$ (4, N=407) = 19.20, p

=0.001. The percentage of SLPs with two or less years since graduation who reported 'lack of skills or confidence with assessing literacy' was significantly higher (26.7%) than SLPs with 6-10 years since graduation (14.5%) and SLPs with 21 or more years since graduation (7.5%) $X^2$ (4, N=407) = 17.53, p =0.002. No significant differences were found in relation to years since graduation and reporting of 'lack of confidence' with assessment for children from CALD backgrounds.

**5.4.5. Sources of information reported by SLPs.** The majority of SLPs (80.6%) indicated 'informal discussion with colleagues' as the most frequent source of information on assessment practices. This was followed by 'formal presentations' (e.g., conferences/workshops) which was selected by 64.1% of SLPs. Less than half of the SLPs surveyed identified 'information provided by employer or professional supervisor' (44.7%), 'journal articles or research reports' (30.0%), 'social media site's (27.3%) or 'online or written material from publishers' (24.3%) as frequent sources of information. The percentage of SLPs who identified each source as a frequent source of information is displayed Figure 5.3.

Group comparisons indicated significant differences between agencies with regards to the frequency with which 'information provided by employer or professional supervisor' and 'social media sites' were selected as main sources of information. 'Information provided by employer or professional supervisor' was significantly more likely to be reported by SLPs in education (57.0%), general agencies (54.5%) and health agencies (45.3%), compared with SLPs in private practice (33.3%) and universities (22.2%) $X^2$ (5, N=407) = 18.27, p =0.003. 'Social media sites' were more likely to be reported by SLPs in private practice (45.7%), compared with SLPs in health agencies (18.8%) and education agencies (12.6%) $X^2$ (5, N=407) = 39.97, p <0.001.
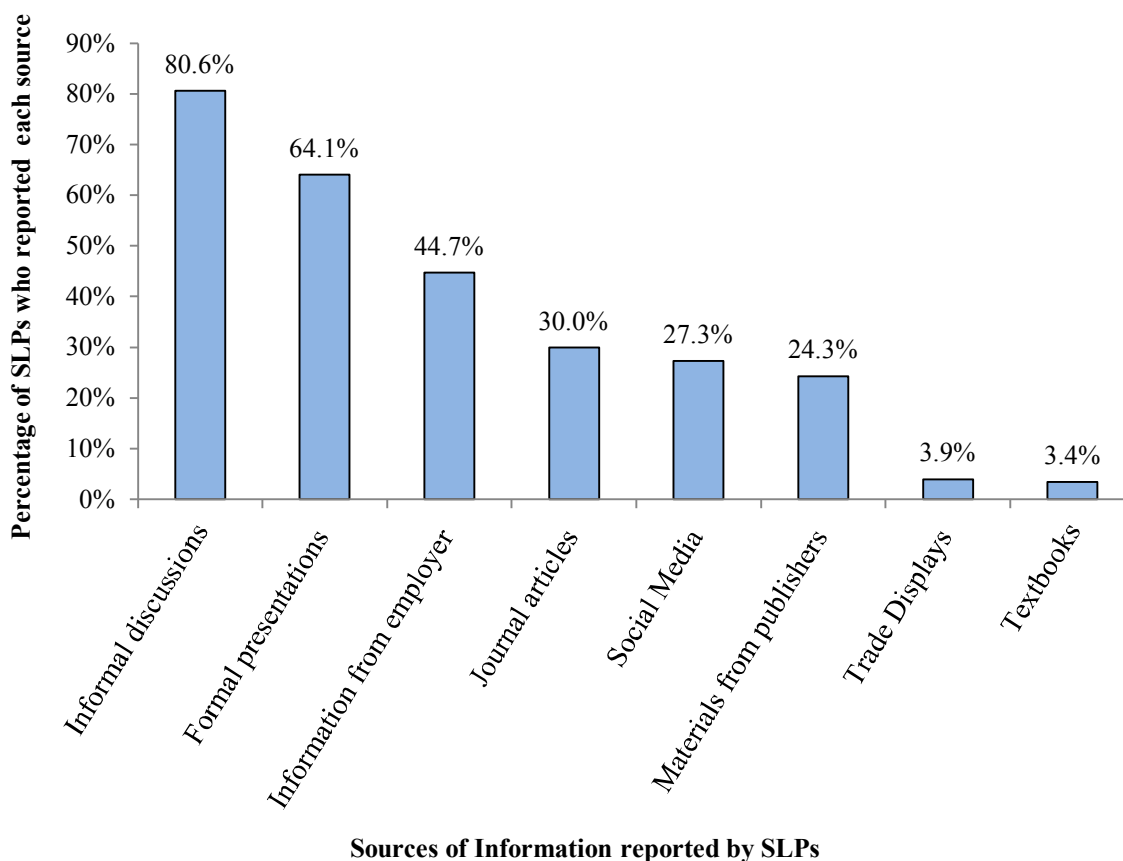
*Figure 5.3.* Percentage of SLPs who identified each source of information on child language assessment (n=407). SLPs were able to select up to three main sources of information.

## 5.5. Discussion

**5.5.1. Regularity of assessment use.** Findings from this survey indicate that most SLPs regularly use assessments that are norm-referenced, de-contextualised and conducted in a clinical context and less regularly use assessments that are contextualised, activity-focused, dynamic or conducted in school or home/community contexts. Given that norm-referenced assessments are typically de-contextualised and conducted in a clinical context, reports of regular use of these three types of assessments is consistent across this survey. These findings are also consistent with findings from previous surveys indicating predominant use of norm-referenced measures by SLPs when assessing the language abilities of children (Beck, 1995; Caesar & Kohler, 2009; Fulcher-Rood et al., 2018; Teoh et al., 2017).

Norm-referenced measures provide important data, however they do have limitations and should be supplemented with information from other types of assessments (Bishop et al., 2016; Trembath, Westerveld, & Shellshear, 2016). Designed specifically to identify if a child's language abilities differ significantly from peers, norm-referenced measures are not suited for selecting intervention goals or measuring intervention outcomes (Huang et al., 1997). Currently available norm-referenced measures also have limitations with regards to diagnostic accuracy, which necessitates a need for data to be collected using a variety of assessments in order to reduce the risk of under-identifying less overt language difficulties (Denman et al., 2017; Harlaar, DeThorne, Smith, Betancourt, & Petrill, 2016). Furthermore, given that norm-referenced measures are only suitable for children whose backgrounds match the normative sample, currently available norm-referenced measures are not appropriate for use with children from CALD backgrounds (Arias & Friberg, 2015; Pearce & Williams, 2013).

It is also important to recognise that scores on de-contextualised assessments may not be representative of a child's ability to participate in daily activities (Thomas-Stonell, Washington, Oddson, Robertson, & Rosenbaum, 2013; Trembath et al., 2016). Numerous studies have identified that children may perform differently on de-contextualised assessments, such as single word vocabulary tests, when compared to contextualised assessments, such as 'language sampling' (Dethorne, Johnson, & Loeb, 2005; Harlaar et al., 2016; Ukrainetz & Blomquist, 2002). Similarly, relationships between language performance in clinical contexts and school or home/community contexts have not always been shown to be concordant, indicating that a child's communicative competence may vary depending on the environmental context in which abilities are assessed (Bishop & McDonald, 2009; Kover, Davidson, Sindberg, & Weismer, 2014). Assessing performance across different tasks and environmental contexts is important for planning services that are functional and relevant to the needs of children and their families (Bishop et al., 2016; Roulstone, 2015; Trembath et

al., 2016). With this in mind, findings from this survey suggest that current SLP practice could be better aligned with current evidence-based practice recommendations by increasing the regularity with which SLPs use child language assessments that are contextualised, activity-focused, dynamic and conducted in school or home/community contexts.

It is acknowledged that multiple factors may act as barriers and facilitators to implementation of evidence-based practice recommendations (Flottorp et al., 2013). By examining the factors that influence child language assessment practice, the challenges SLPs report and sources from which SLPs gain information on child language assessment practice; a greater understanding of the contextual factors that influence practice may be gained (Olswang & Prelock, 2015). This in turn may assist in identifying actions that may facilitate successful practice change.

**5.5.2. Factors that influence assessment use.** This study identified that service agency may influence SLP language assessment practice. The finding that SLPs in private practice had at least three times greater odds than SLPs in disability agencies of reporting regular use of norm-referenced, de-contextualised and clinical context assessments likely reflects that existing norm-referenced measures are less suitable for children with severe disorders and are thus used less for this population of children (Fulcher-Rood et al., 2018; Watson & Pennington, 2015). However, this finding may also reflect that SLPs in different agencies take a different focus when collecting assessment data.

Interestingly, the odds of SLPs in disability and education agencies reporting regular use of activity-focused assessments were 15 and 4 times greater respectively than odds than SLPs in private practice. SLPs in disability and education agencies also had greater odds than SLPs in private practice of reporting frequent use of school context assessments, with SLPs in private practice in turn having greater odds than SLPs in health agencies of reporting frequent use of school context assessments. A difference in education and disability agencies relative to private practice and health agencies may be workplace requirements for school

performance or functional abilities to be explicitly targeted in assessment, thus necessitating use of assessments that are activity-focused and directed specifically towards a school context. This is also consistent with results from this study indicating that SLPs in disability and education agencies were also more likely to identify workplace requirements (i.e. workplace requires particular data or use of particular tools) as a main challenge. It is possible that SLPs predominantly choose to use de-contextualised, norm-referenced measures unless their workplace specifically requires them to collect data on functional performance.

The odds of SLPs reporting frequent use of contextualised and dynamic-gradual prompting assessments was greater for SLPs with more years since graduation. Unlike de-contextualised assessments, few contextualised or dynamic assessments with set guidelines for administration and scoring exist; therefore this finding may reflect that SLPs with more years of experience have more confidence and skill conducting assessment procedures without set guidelines for administration and scoring and using their clinical experience to interpret the data gathered from these assessments (K. S. Wilson et al., 1991).

Use of contextualised and both types of dynamic assessments were also influenced by Australian State, with SLPs in Western Australia having at least two times greater odds than SLPs in New South Wales of reporting frequent use of these assessments. Previous literature has identified that the popularity of specific language measures may vary regionally (Westerveld & Claessen, 2014) and it is possible that a similar tendency occurs with regards to the types of assessments SLPs use. Differences between states could be attributable to policy influences that span across states, rather than agencies. For example, in Western Australia, a state-wide process exists for accessing specialised schooling, thus creating a situation where SLPs from different agencies in Western Australia are required to use the same types of assessments when completing school applications (North East Metropolitan Language Development Centre & Outreach Service). It is also possible that differences between university training programs contribute to variations across states.

Another finding from this study was that SLPs with a higher proportion of children from CALD on their caseloads did not report less regular use of norm-referenced measures, despite these assessments being less suitable for this population of children (Pearce & Williams, 2013; Teoh, Brebner, & McCormack, 2012). Similarly, although dynamic assessments can be suitable alternatives for children from CALD backgrounds (Kapantzoglou, Restrepo, & Thompson, 2012; Peña, Gillam, & Bedore, 2014), these same SLPs were not more likely to regularly use dynamic assessments. The inappropriate use of norm-referenced measures with children from CALD backgrounds has been reported in previous literature (Arias & Friberg, 2015; Caesar & Kohler, 2009; Teoh et al., 2017) and represents an area of child language assessment practice that requires significant change.

Factors that influence the regularity with which criterion-referenced/descriptive and home/community context assessments were used were not identified in this study. This indicates that, unlike other types of assessment practices, regular use of these types of assessments is influenced less by contextual factors such as service agency, geographical location and SLP experience.

**5.5.3. Challenges reported by SLPs**. In this survey, the most frequently reported challenges by SLPs related to lack of time. This included 'limited time to plan and analyse assessment', 'limited face-face time with children' and 'limited time to meet with teachers and parents'. This challenge was also identified from previous surveys, with 'limited time' reported as a reason for the over-reliance of norm-referenced measures and as a barrier for use of contextualised and dynamic assessments (Arias & Friberg, 2015; Fulcher-Rood et al., 2018; Huang et al., 1997; Pavelko et al., 2016; Westerveld & Claessen, 2014). Norm-referenced language measures have set guidelines for administration and scoring makes these assessments easy to administer and interpret (Fulcher-Rood et al., 2018). In contrast, activity-focused and dynamic assessments likely require more time to plan, administer and interpret results. By their very nature, assessments that are activity-focused, dynamic or conducted in

school or, home/community contexts may always require some individualisation for a child's particular activities or contexts. Nonetheless, the development of more specific guidelines for administering and interpreting these types of assessments may assist in addressing some of the time barriers SLPs experience.

'Lack of assessment materials' (e.g. due to budget constraints) was also reported as a main challenge by SLPs. The finding that this challenge was less frequently reported by SLPs in education agencies and universities may be reflective of greater financial capacity to purchase resources in such larger organisations. However, the types of assessments that SLPs reported as using less regularly, such as activity-focused and dynamic assessments, do not typically require high material resourcing or financial outlay. For this reason, it seems unlikely that lack of assessment materials specifically influences the types of assessments SLPs regularly use, although further research is needed to examine this. Instead, it is possible that budget constraints influence the particular norm-referenced measures SLPs choose to purchase, for example, cost may influence decision-making when choosing between two norm-referenced measures.

SLPs also reported challenges related to 'limited training in assessments or lack of skills or confidence' with assessing specific populations, such as children from CALD backgrounds, children with complex communication needs or children with literacy difficulties. Complex communication needs and literacy have been identified in previous surveys as areas in which SLPs may lack confidence (Balandin & Iacono, 1998; Blood, Mamett, Gordon, & Blood, 2010). As it likely takes time to develop confidence and skill in these more complex areas of professional practice, the finding in this survey that SLPs with fewer years since graduation are more likely to report lack of skills or confidence with assessing complex communication needs or literacy is not surprising. Assessment for children from CALD backgrounds was the most frequently identified challenge related to 'lack of skills and confidence' and was not identified as being related to years since graduation. This

finding reflects a priority for professional development across the profession, particularly given the identified over-reliance on norm-referenced measures for this population of children (Arias & Friberg, 2015; Teoh et al., 2017). In particular, developing SLPs skills in conducting dynamic assessments may help facilitate evidence-based diagnostic assessment practices for children from CALD backgrounds (Teoh et al., 2017)

The physical location in which services are provided may also contribute to variations in use of different types of assessments. SLPs in education agencies are more likely to be located on school grounds and may be more easily able to meet with teachers or visit classrooms to observe children in daily school environments (Koole, Nelson, & Curtis, 2015). This appeared to be reflected in the survey findings, with SLPs working in education agencies being less likely to report 'setting constraints' compared to SLPs in other agencies. Interestingly, SLPs in education agencies were more likely to report 'lack of time to meet with teachers' as a challenge, despite being the group most likely to be able access teachers easily. It is possible that physical location is initially perceived as the greatest challenge, and when this challenge is removed, other challenges emerge, such as being able to schedule time to collect information from teachers.

**5.5.4. Sources of information reported by SLPs.** In this survey, the most frequently reported source of information was 'informal discussion with colleagues'. This is consistent with findings from previous studies indicating that SLPs tend to obtain information on clinical practice from colleagues, workplaces or workshops rather than research articles (Pavelko et al., 2016; Singh et al., 2016; K. S. Wilson et al., 1991). It is possible that this tendency contributes to variations across states or agencies as clusters of SLPs may develop similar practices by sharing information amongst each other and it is speculated that this may be a reason for differences in assessment practice between states in this study. Given that 'informal discussions with colleagues' is the most frequently reported source of information,

it is important to ensure that this tendency is harnessed to promote sharing of evidence-based information.

The finding that SLPs in education and health agencies are more likely to report 'information from employer or professional supervisor' as a source of information is likely attributable to these services being most typically provided through large government organisations that have a greater structure for professional supervision compared with smaller agencies and private practices. Nonetheless, it raises the possibility that on-the-job training may contribute to variations in practice across agencies. For example, previous studies suggested that SLPs may not graduate well-equipped to work within contemporary educational service delivery models and that information provided in workplaces may be a primary source of information on educational service provision for SLPs working in educational settings (Sanger, Snow, Colburn, Gergen, & Ruf, 2012; L. Wilson, McNeill, & Gillon, 2015). Additionally, having access to experts in workplaces appears to have a positive influence on SLP practice (Koole et al., 2015). Therefore, it is possible that more frequent use of assessments that target a school context by SLPs in education agencies is influenced by the professional development opportunities provided in workplaces, such as training in assessing children's ability to participate in the school curriculum.

**5.5.5. Implications**. Evidence-based practice recommends that SLPs should use a range of different types of assessments when assessing the language abilities of primary school children (Bishop et al., 2016). However, findings from this survey indicate that current clinical practice may be aligned with this recommendation as SLPs predominantly use of assessments that are norm-referenced, de-contextualised and conducted in a clinical context; and less regularly use other types of assessments. When contextualised and activity-focused language assessments are not routinely conducted, there is risk that language difficulties at a discourse level will be under-identified and that functional performance in daily activities will not be adequately assessed. Infrequent use of dynamic assessments and assessments that

target school and home/community contexts may to lead intervention goals and classroom supports that are not well-matched to functional needs. Similarly, the use of norm-referenced measures rather than use of dynamic assessments with children from CALD backgrounds may lead to inappropriate conclusions being drawn regarding the language abilities of these children.

For these reasons, actions that increase the frequency with which SLPs use assessments that are contextualised, activity-focused, dynamic and targeted at daily environments are necessary to advance clinical practice in the field of child language assessment. As variations exist between service agencies with regards to the regularity with which different assessments were used and the frequency with which different challenges are reported, it is important that actions to advance clinical practice are developed with consideration of the unique contexts within different service agencies.

**5.5.6. Future directions**. Findings from this study suggest a number of future directions that may facilitate implementation of evidence-based child language assessment practices by SLPs. This includes the creation of activity-focused and dynamic assessments that have set guidelines for administration and analysis. There is also a need for questionnaires and interview protocols that facilitate collection of data on a child's abilities specific school or home/community contexts.

As SLPs report challenges related to limited time, focus needs to be placed on optimising the balance between the time needed to conduct quality assessment and costs of professional time, with this reflected in workload and funding policies. Nonetheless, it is also important to build SLP capacity to conduct assessment in time efficient manner, for example additional training in administering and analysing different types of assessments may lead to SLPs being more time efficient when conducting assessment.

To build SLP capacity, university programs should be examined to ensure that entry-level SLPs are sufficiently prepared to conduct different types of assessments (Pavelko et al.,

2016). It is also important that SLPs have access to continuing professional development post-graduation through workshops and consultation with SLPs who have expert knowledge in child language assessment. Given that less than half of SLPs report journal articles as a frequent source of information on assessment practices, actions that increase the accessibility and utilisation of journal articles by SLPs may also be needed to support implementation of evidence-based practice recommendations (Reilly, 2004).

There is also a need for future research examining child language assessment practice. For example, further research is needed to examine the language assessment procedures used by SLPs in different countries. Although SLP practice across English-speaking countries is likely informed by the same international literature, factors external to individual SLPs may result in differences between countries (Singh et al., 2016). Future research would ideally use the same assessment definitions and response scales to allow findings from different countries to be compared with consistency and transparency.

Finally, since this study identified variations with regards to the frequency with which different groups of SLPs use different types of assessments, future investigations are needed to examine the implications of these differences for children and their families. Given that intervention is planned based on assessment findings, it likely that differences in assessments across agencies or states are associated with differences in interventions provided; however further research is needed to examine this (Roulstone, 2001).

**5.5.7. Limitations.** The survey used in this study utilised terminology for describing different types of assessments from a newly developed taxonomy that was agreed-upon through a previous Delphi technique involving Australian SLPs with expertise in child language. This is an important methodological advance to previous survey research in the field of child language as it ensured greater consistency across survey participants with regards to how assessment practices were described.

A limitation of this study is that, although the study included a large sample of SLPs from all over Australia; some groups, such as SLPs in smaller states or agencies, had small sample sizes. This limits the extent that survey findings can be generalised for these groups of SLPs. As some of these groups have small overall populations (for example, the 11 SLPs from Northern Territory in the current study represent 40.1% of the estimated population of SLPs who work with primary school-aged children in the Northern Territory), qualitative methodologies, such as semi-structured interviews, may be more appropriate for understanding the types of assessments used in these more unique contexts.

This study specifically examined language assessment practices for primary school children. SLPs may not use the same types of assessments with the same regularity when assessing children of other ages (Caesar & Kohler, 2009; Pavelko et al., 2016). Similarly, this survey did not examine all possible factors that may influence the regularity with which SLPs may use different types of assessments, such as pre-existing diagnoses or severity of language disorder. This study also did not examine the processes SLPs employ when analysing data collected from different types of assessments. Therefore, further investigation is needed to examine these areas of SLP assessment practice.

## 5.6. Conclusion

SLPs predominantly use norm-referenced and de-contextualised assessments conducted in clinical context when assessing the language abilities of primary school children and less regularly use contextualised, activity-focused or dynamic assessments and assessments conducted in everyday environmental contexts. Factors that influence the use of different types of assessments were identified as service agency, Australian State and years since SLP graduation. SLPs identified challenges related to 'limited time', 'lack of assessment materials', 'limited access to training in assessment' and 'lack of skill or confidence with assessing children from CALD backgrounds'. The most frequently reported source of information on assessments was 'informal discussions with colleagues'. Given

current recommendations for practice, future development in the field of child language

assessment should focus on actions that increase the regularity with which contextualised,

activity-focused and dynamic assessments are used and the regularity with which SLPs assess

abilities in school or home/community contexts. These actions should be developed with

consideration to unique contextual differences between service agencies and states and the

challenges that SLPs report in relation to language assessment for primary school children.

**References for Chapter 5**

Andrews, D., Nonnecke, B., & Preece, J. (2003). Electronic survey methodology: A case study in reaching hard to involve Internet Users. *International Journal of Human-Computer Interaction, 16*(2), 185-210.

Arias, G., & Friberg, J. (2015). Bilingual language assessment: Contemporary versus recommended practice in American schools. *Language, Speech, and Hearing Services in Schools, 48*(1), 1-15. doi:0.1044/2016_LSHSS-15-0090

Australian Bureau of Statistics. (2016). Australian Statistical Geography Standard (ASGS). Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.005

Balandin, S., & Iacono, T. (1998). AAC and Australian speech pathologists: Report on a national survey. *Augmentative and Alternative Communication, 14*(4), 239-249. doi:10.1080/07434619812331278416

Beck, A. R. (1995). Language assessment methods for three age groups of children. *Journal of Communication Development, 17*(2), 51-56. doi:10.1177/152574019501700206

Betz, S. K., Eickhoff, J. K., & Sullivan, S. F. (2013). Factors influencing the selection of standardized tests for the diagnosis of Specific Language Impairment. *Language, Speech and Hearing Services in Schools, 44*, 133-143. doi:10.1044/0161-1461(2012/12-0093)

Bishop, D. V. M., & McDonald, D. (2009). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of Language & Communication Disorders, 44*(5), 600-615. doi:10.1080/13682820802259662

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS ONE, 11*(7), e0158753. doi:10.1371/journal.pone.0158753

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE-2 consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*. doi:10.1111/jcpp.12721

Blais, J. G., & Grondin, J. (2011). The influence of labels associated with anchor points of likert-type response scales in survey questionnaires. *Journal of Applied Measurement, 12*, 370–386.

Blood, G. W., Mamett, C., Gordon, R., & Blood, I. M. (2010). Written language disorders: Speech-language pathologists' training, knowledge, and confidence. *Language, Speech, and Hearing Services in Schools, 41*(4), 416-428. doi:10.1044/0161-1461(2009/09-0032)

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*, 190-200. doi:10.1044/0161-1461(2007/020)

Caesar, L. G., & Kohler, P. D. (2009). Tools clinicians use: A survey of language assessment procedures used by school-based speech-pathologists. *Communication Disorders Quarterly, 30*(4), 226-236. doi:10.1177/1525740108326334

Denman, D., Cordier, R., Munro, N., Kim, J.-H., & Speyer, R. (Under Review). Language Assessment for Primary School Children (Part II): What Reasons Drive SLP Choice? .

Denman, D., Kim, J.-H., Munro, N., Speyer, R., & Cordier, R. (2019). Describing language assessments for school-aged children: A Delphi study. *International Journal of Speech Language Pathology*, Advance Online Publication. doi:10.1080/17549507.2018.1552716

Denman, D., Speyer, R., Munro, N., Pearce, W., Chen, Y., & Cordier, R. (2017).

    Psychometric properties of language assessments for children aged 4-12 years: A

    systematic review. *Frontiers in Psychology, 8*, 1-28. doi:10.3389/fpsyg.2017.01515

Dethorne, L. S., Johnson, B. W., & Loeb, J. W. (2005). A closer look at MLU: What does it

    really measure? *Clinical Linguistics & Phonetics, 19*(8), 635-648.

    doi:10.1080/02699200410001716165

Eadie, P. (2003). Speech pathology assessment practices: One assessment or many? *Advances*

    *in Speech Language Pathology, 5*(1), 65-68. doi:10.1080/14417040510001669081

Flottorp, S. A., Oxman, A. D., Krause, J., Musila, N. R., Wensing, M., Godycki-Cwirko, M.,

    . . . Eccles, M. P. (2013). A checklist for identifying determinants of practice: A

    systematic review and synthesis of frameworks and taxonomies of factors that prevent

    or enable improvements in healthcare professional practice. *Implementation Science,*

    *8*(1), 1-11. doi:10.1186/1748-5908-8-35

Fulcher-Rood, K., Castilla-Earls, A. P., & Higginbotham, J. (2018). School-based speech-

    language pathologists' perspectives on diagnostic decision making. *American Journal*

    *of Speech-Language Pathology, 27*(2), 796-812. doi:10.1044/2018_AJSLP-16-0121

Harlaar, N., DeThorne, L. S., Smith, J. M., Betancourt, M. A., & Petrill, S. A. (2016).

    Longitudinal effects on early adolescent language: A twin study. *Journal of Speech,*

    *Language, and Hearing Research, 59*(5), 1059-1073. doi:10.1044/2016_JSLHR-L-

    15-0257

Hoffman, L. M., Ireland, M., Hall-Mills, S., & Flynn, P. (2013). Evidence-based speech-

    language pathology practices in schools: Findings from a national survey. *Language,*

    *Speech, and Hearing Services in Schools, 44*, 266-280. doi:10.1044/0161-

    1461(2013/12-0041)

Huang, R. J., Hopkins, J., & Nippold, M. A. (1997). Satisfaction with standardized language testing: A survey of speech-language pathologists. *Language, Speech, and Hearing Services in Schools, 28*(1), 12-29. doi:10.1044/0161-1461.2801.12

Hux, K., Morris-Friehe, M., & Sanger, D. D. (1993). Language Sampling Practices: A Survey of Nine States. *Language, Speech, and Hearing Services in Schools, 24*, 84-91. doi:10.1044/0161-1461.2402.84

IBM Corp. (Released 2011). IBM SPSS Statistics for Windows Version 20.0. Armonk, NY: IBM Corp.

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)

Kemp, K., & Klee, T. (1997). Clinical language sampling practices: Results of a survey of speech-language pathologists in the United States. *Child Language Teaching and Therapy, 13*(2), 161-176. doi:10.1177/026565909701300204

Koole, H., Nelson, N. W., & Curtis, A. B. (2015). Factors influencing choices of contextualized versus traditional practices with children and adolescents who have traumatic brain injury. *Language, Speech, and Hearing Services in Schools, 46*, 352-361. doi:10.1044/2015_LSHSS-14-0109

Kover, S. T., Davidson, M. M., Sindberg, H. A., & Weismer, S. E. (2014). Use of the ADOS for assessing spontaneous expressive language in young children with ASD: A comparison of sampling contexts. *Journal of Speech, Language, and Hearing Research, 57*(6), 2221-2233. doi:10.1044/2014_JSLHR-L-13-0330

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., . . . Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of

language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry, 57*(11), 1247-1257. doi:10.1111/jcpp.12573

North East Metropolitan Language Development Centre & Outreach Service. Language Development Centres. Retrieved from http://northeastldc.wa.edu.au/wp-content/uploads/2017/05/NEMLDC-SCHOOL-BROCHURE.pdf

Olswang, L. B., & Prelock, P. A. (2015). Bridging the gap between research and practice: Implementation science. *Journal of Speech, Language, and Hearing Research, 58*, 1818-1826. doi:10.1044/2015_JSLHR-L-14-0305

Oppenheimer, A. J., Pannucci, C. J., Kasten, S. J., & Haase, S. C. (2011). Survey says? A primer on web-based survey design and distribution. *Plastic and Reconstructive surgery, 128*(1), 299-304. doi:10.1097/PRS.0b013e3182174413

Pavelko, S. L., Owens, R. E., Ireland, M., & Hahs-Vaughn, D. L. (2016). Use of language sample analysis by school-based SLPs: Results of a nationwide survey. *Language, Speech, and Hearing Services in Schools, 47*(3), 246-258. doi:10.1044/2016_LSHSS-15-0044

Pearce, W. M., & Williams, C. (2013). The cultural appropriateness and diagnostic usefulness of standardized language assessments for Indigenous Australian children. *International Journal of Speech-Language Pathology, 15*(4), 429-440. doi:10.3109/17549507.2012.762043

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

Qualtrics. (2005). Qualtrics March 2018 Edition. Provo, Utah, USA: Qualtrics. Retrieved from https://www.qualtrics.com

Reilly, S. (2004). The challenges in making speech pathology practice evidence based. *Advances in Speech–Language Pathology, 6*(2), 113-124. doi:10.1080/14417040410001708549

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

Rosenfeld, M. (2002). *Report on the ASHA speech-language pathology health care survey*. Retrieved from Rockville, MD:

Roulstone, S. (2001). Consensus and variation between speech and language therapists in the assessment and selection of preschool children for intervention: A body of knowledge or idiosyncratic decisions? *International Journal of Language and Communication Disorders, 36*(3), 329–348. doi:10.1080/13682820010019928

Roulstone, S. (2015). Exploring the relationship between client perspectives, clinical expertise and research evidence. *International Journal of SpeechLanguage Pathology, 17*(3), 211-222. doi:10.3109/17549507.2015.1016112

Roulstone, S., Marshall, J., Powell, G. G., Goldbart, J., Wren, Y. E., Coad, J., . . . Coad, R. A. (2015). Evidence-based intervention for preschool children with primary speech and language impairments: Child Talk – an exploratory mixed-methods study. *Programme Grants for Applied Research, 3*(5), 1-408. doi:10.3310/pgfar03050

Ruggero, L., McCabe, P., Ballard, K. J., & Munro, N. (2012). Paediatric speech-language pathology service delivery: An exploratory survey of Australian parents. *International Journal of Speech-Language Pathology, 14*(4), 338-350. doi:10.3109/17549507.2012.650213

Sanger, D., Snow, P. C., Colburn, C., Gergen, M., & Ruf, M. (2012). Speech-language pathologists' reactions to response to intervention: A qualitative study. *International*

*Journal of Speech-Language Pathology, 14*(1), 1-10.
doi:10.3109/17549507.2011.604793

Singh, J. S., Chan, M. Y., & Rusli, Y. A. (2016). Practise patterns of Malaysian speech-
language pathologists in managing children with speech and language delay/disorder.
*International Journal of Speech-Language Pathology, 18*(6), 560-570.
doi:10.3109/17549507.2016.1139624

Spaulding, T. J., Szulga, M. S., & Figueroa, C. (2012). Using norm-referenced tests to
determine severity of language impairment in children: Disconnect between US
policy makers and test developers. *Language, Speech, and Hearing Services in
Schools, 43*(2), 176-190. doi:10.1044/0161-1461(2011/10-0103)

Speech Pathology Australia. (2014). *Submission to the inquiry into the prevalence of different
types of speech, language and communication disorders and speech pathology
services in Australia.* Retrieved from Australia:
https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Aff
airs/Speech_Pathology

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia medica, 24*(1),
12-18. doi:10.11613/BM.2014.003

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, 21*(1), 10-21.
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:10.3109/17549507.2011.65217

Thomas-Stonell, N., Washington, K., Oddson, B., Robertson, B., & Rosenbaum, P. (2013). Measuring communicative participation using the FOCUS©: Focus on the Outcomes of Communication Under Six. *Child: care, health and development, 39*(4), 474-480. doi:10.1111/cch.12049

Trembath, D., Westerveld, M., & Shellshear, L. (2016). Assessing spoken language outcomes in children with ASD: A systematic review. *Current developmental Disorders Reports, 3*(33), 33-45. doi:10.1007/s40474-016-0068-8

Ukrainetz, T. A., & Blomquist, C. (2002). The criterion validity of four vocabulary tests compared with a language sample. *Child Language Teaching and Therapy, 18*(1), 59-78. doi:10.1191/0265659002ct227oa

Watson, R. M., & Pennington, L. (2015). Assessment and management of the communication difficulties of children with cerebral palsy: A UK survey of SLT practice. *International Journal of Language & Communication Disorders, 50*(2), 241-259. doi:10.1111/1460-6984.12138

Westerveld, M. F., & Claessen, M. (2014). Clinician survey of language sampling practices in Australia. *International Journal of Speech-Language Pathology, 16*(3), 242-249. doi:10.3109/17549507.2013.871336

Williams, C., & McLeod, S. (2012). Speech-language pathologists assessment and intervention practices with multi-lingual children. *International Journal of Speech-Language Pathology, 14*(3), 292-305. doi:10.3109/17549507.2011.636071

Wilson, K. S., Blackmon, R. C., Hall, R. E., & Elcholtz, G. E. (1991). Methods of language assessment: A survey of California public school clinicians. *Language, Speech and Hearing Services in Schools, 22*, 236-241. doi:10.1044/0161-1461.2204.236

Wilson, L., McNeill, B., & Gillon, G. T. (2015). The knowledge and perceptions of prospective teachers and speech language therapists in collaborative language and

literacy instruction. *Autism & Developmental Language Impairments, 2*, 1-17.

doi:10.1177/0265659015585374

**Supplementary Appendix 5.1.**

Survey for Australian Speech-Language Pathologists (Questions Relating to Part I)

**I consent to answering questions in an on-line survey and for my responses to be used**

**for the purposes described above**

*[Yes/No response. If no, skip to end of survey]*

**SECTION 1**

**Do you have current practicing membership with Speech Pathology Australia?**

*[Yes/No response]*

**Please indicate your age:**

*[Multiple choice response]*

**Please indicate your gender:**

*[Multiple choice response]*

**Is English your first language?**

*[Yes/No response]*

**Please indicate the number of years since you graduated as a speech pathologist:**

*[Multiple choice response]*

**Are you currently in paid employment as an SLP?**

(This may include non-clinical SLP roles*)*

*[Yes/No response. If no, skip to end of survey]*

**Please give your postcode**

If you are currently employed please give your work postcode.

If you work in more than one job, please give the postcode the job in which you work the

most hours or if hours are equal, choose the job in which you have worked the longest.

If you are not currently employed, please give your home postcode (e.g. full-time student or

are on extended leave)

*[Open text response]*

**Please indicate the amount of time per week you are currently employed (paid to work) as a speech pathologist (across all jobs):**

*[Multiple choice response]*

**Please select the box/es that best describe/s your current paid work as a speech pathologist** (**across all jobs**).

Select all appropriate options e.g. if you provide services to children aged 0-6 years then select both of the first two options; or if you provide services to clients of all ages then select all the clinical service provision boxes.

*[Multiple choice response. If participant does not select response "work clinically with children 4-12 years including supervision of students providing clinical services" then skip to end of survey]*

**In the last 12 months, have you provided clinical services to at least 40 children aged 4-12 years with oral or written language disorders?** (Includes supervision of SLP students who provide clinical services).

For the purposes of this survey, the term 'children with language disorder' is used broadly to refer to any children who require support for oral or written language (i.e. semantics, syntax, morphology, phonemic awareness, discourse or social abilities), regardless of the primary diagnosis, aetiology or co morbidities associated with the language support needs.

Children who have a lack of familiarity with standard Australian English are also included in this group if they are accessing your services for language assessment.

This includes children with: developmental language disorder, dyslexia, autism spectrum disorder, learning difficulties, intellectual disability or language disorder associated with

conditions such as traumatic brain injury or hearing impairment.

*[Yes/No response. If no, skip to end of survey]*

**SECTION 2**

**Indicate the option that best describes the agency through which you provide clinical services to children aged 4-12 years with language disorder:**

If you work in more than one job, please complete this survey for the job through which you most frequently service children aged 4-12 years with language disorder.

*[Multiple choice response]*

**Consider the last 40 children (4-12 years with language disorder) who accessed your services. How many children <u>did not</u> have standard Australian English as their first or only language?** (i.e. how many children were bi-lingual, learning English as a second language or were considered as having a cultural/linguistic difference)

*[Multiple choice response]*

**Please indicate if any of the following service criteria or eligibility requirements exist in the agency where you work in order for children to receive language assessment.**

You may select more than one option.

*[Multiple choice responses, with open text boxes for responses not listed]*

**Indicate the most frequent sources from which you obtain information about assessment tools or procedures or learn about new assessments.**

You may choose up to three main sources of information.

*[Multiple choice responses with open text boxes for responses not listed]*

**Indicate if you perceive any of the following as challenges in your workplace with regards to language assessment.**

You may choose up to four main challenges.

*[Multiple choice responses with open text boxes for responses not listed]*

**SECTION 3**

**In this section of the survey you are asked to estimate the frequency in which different types of oral and written assessments were used** (considering the last 40 children who accessed your services for assessment).

Terminology may be used differently across the profession; therefore, it is important that you read the survey information carefully and select answers based on how terms are defined in this survey (and not how you define them or have seen them defined elsewhere).

Please open the following link to see examples of how commonly used assessments are categorised in the survey: Table of Categorised Assessments

When answering the questions, please consider:

Assessments conducted by yourself as well as any assessments that you or your service agency supported/trained others to conduct on your behalf

All assessment data gathering activities (e.g. tests, observations, parent or teacher interviews and collection of case histories)

Assessments conducted for any purpose (e.g. diagnosis, screening, detecting change, selecting intervention**)**

*[Likert scale response]*

**Norm-Referenced Assessments**

A student's performance is compared to the performance of a sample of aged matched peers (i.e. assessment provides a standard score, percentile rank, age-equivalent score or mean and standard deviation for a sample population).

Note: Norm-referenced tests that are not scored using norms are not considered norm-referenced

Examples include: * CELF-4 Core Language Index Subtests  * CELF-5 Reading Comprehension and Writing Comprehension or Structured Writing Tests  * CELF-4

Screening Test  * PLS-5  *  CASL  * Renfrew Action Picture Test  *  TNL or TNL-

2  * TOPL-2  * TOPS  *  CCC-2

For further examples please see the Table of Categorised Assessments

**How often were norm-referenced tests used as part of a child's assessment?** (Considering

the last 40 children who were assessed)

*[Likert scale response]*

## Criterion Referenced or Descriptive Assessments

These assessments do not provide norms for a sample of aged matched peers, but are used to

compare a child's performance against a pre-determined level or expected criterion (e.g.

developmental expectations, curriculum level or cut score) or are only used to describe

performance qualitatively.

Note: All assessments that are not norm-referenced must be criterion referenced/descriptive.

Examples include: * Norm-referenced assessments that are altered and not scored using

norms * Assessments that do not provide standard scores, percentile ranks or means and

standard deviations from an age matched sample * Tests that you designed yourself

 * Questionnaires or interviews that do not follow a set protocol

For further examples see the Table of Categorised Assessments

**How often were criterion-referenced or descriptive assessments used as part of a child's**

**assessment?** (Considering the last 40 children who were assessed)

*[Likert scale response]*

## Discrete Skill Tests (De-Contextualised)

Discrete skills are assessed in a short-answer 'test format' that is highly directed by the

SLP (e.g. picture naming tasks, picture description tasks, sentence production tasks, defining

words, word or sentence repetition tasks, identifying social expectations, identifying and

segmenting sounds in words, spelling individual words)

Examples include: * CASL  * CELF-5 Core Language Index subtests * CELF-5 Pragmatics Profile  * CELF-4 Screening Test  * CCC-2  * Communication Matrix  * PLS-5  * Renfrew Action Picture Test  * SPAT-R  * TOPL-2  * TOPS-3  * Short answer 'tests' that you create yourself. For further examples see the Table of Categorised Assessments.

**How many children were assessed using discrete-skill tests?** (Considering the last 40 children who were assessed)

*[Likert scale response]*

**Discourse or Text Level Assessments (Contextualised)**

Assessment tasks are directed by the SLP, but occur in meaningful communicative situations (e.g. oral text comprehension, written text comprehension, narrative telling, story or report writing, language sampling during conversation or play)

Note: Assessments that are not discrete skill tests must be either discourse/text level or an assessment in daily activities (see next question).

Examples include: * CELF-5 Pragmatic Activities Checklist  * Renfrew Bus Story  * Peter and the Cat Narrative Assessment  * TNL  * Neale Analysis of Reading Ability  * CELF-5 Reading Comprehension and Structured Writing Tests  * OWLS-II Reading and Writing Tests  *  Written language sampling tasks that you designed yourself  *  Oral language sampling tasks that you designed yourself  *  Role play tasks designed to represent situations that are relevant to child's daily life. For further examples please see the Table of Categorised Assessments.

**How many children were assessed at text or discourse level?** (Considering the last 40 children who were assessed)

*[Likert scale response]*

**Assessment in Daily Activities (Activity-Focused)**

The child's abilities are assessed within the actual activities in which they participate, with the

assessor observing or being part of the interaction rather than directing the tasks.

Note: Assessments that are not discrete skill tests must be either discourse/text level (see previous question) or an assessment in daily activities.

Examples Include: * Assisting a child to complete a class assignment and assessing the level of support they require * Observing a child during a regular class lesson and noting abilities or behaviours * Observing a child during free play with peers at lunchtime and noting abilities and behaviours * Observing the child whilst ordering food from a cafe and noting abilities and behaviours. For further examples please see the Table of Categorised Assessments.

**How many children were assessed during daily activities?** (Considering the last 40 children who were assessed)

*[Likert scale response]*

## Assessments Conducted in a Clinical Context

These assessments are conducted away from the child's regular environment (i.e. conducted in a clinic room or a withdrawal room at school) using clinical materials (e.g. a standardised norm-referenced test)

Examples include: * CELF-4 Core and Language Index Tests * CASL * PLS-5 * TNL * ERRNI * SALT * NEALE * Observations by the SLP of the child interacting in a clinical environment with clinical materials (e.g. during an SLP session conducted in a consultation room or withdrawal room). For further examples see the Table of Categorised Assessments.

**How many children were assessed in a clinical context?** (Considering the last 40 children who were assessed

*[Likert scale response]*

**<u>Assessments Conducted in a School Context</u>**

Skills are assessed with specific reference to a school context

<u>Note:</u> Assessments that are not clinical context must have either a school or home/community

context (see next question)

Examples include: * Information reported by a teacher regarding communication at school

(e.g. CCC-2, Communication Matrix, CELF-4 Pragmatics Profile, teacher interview)  *

Assessments conducted by the SLP using materials from the curriculum (e.g. vocabulary is

assessed using words from the current curriculum unit of work; Reading Comprehension is

assessed using the same book that is being studied in class)  * Observations by the SLP

during class or during activities similar to class activities (e.g. assisting the child to complete

a class assignment and noting level of support required * Observing communication skills in

class). For further examples please see the <u>Table of Categorised Assessments.</u>

**How many children were assessed in a school context?** (Considering the last 40 children

who were assessed)

*[Likert scale response]*

**<u>Assessments Conducted in a Home/Community Context</u>**

Skills are assessed with reference to a home/community context

<u>Note:</u> Assessments that do not have a clinical context must have either a school (see previous

question) or home/community context.

Examples include: * Information reported by a parent or caregiver regarding communication

at home (e.g. CCC-2, Communication Matrix, CELF-4 Pragmatics profile, parent/caregiver

interview)  * Observations conducted by an SLP during usual daily activities - not in the

classroom (e.g. ordering food from a cafe, free play with peers or siblings). For further

examples please see the <u>Table of Categorised Assessments.</u>

**How many children were assessed in a home or community context?** (Considering the last 40 children who were assessed)

*[Likert scale response]*

## Test-Teach-Retest (Dynamic) Assessments

Investigation of a child's learning potential by testing skills, teaching skills that child does not know and then retesting to determine response to teaching. The testing - teaching – retesting process occurs across one or two assessment sessions.

Examples include: * Testing vocabulary in a picture naming task, teaching words child did not know and then testing again to determine response to teaching * Collecting a narrative retell, explaining/teaching features of a good narrative story and then collecting another narrative retell

**How many children were assessed using test-teach-retest procedures?** (Considering the last 40 children who were assessed)

*[Likert scale response]*

## Gradual Prompting or Task Modification (Dynamic) Assessments

Investigation of a child's learning potential or the level of support that the child requires by assessing skills under different conditions or with varied levels of prompting/scaffolding.

Examples include: * Assessing a child's narrative retell skills with and without visual picture supports and comparing performance * Assessing a child's ability to answer questions using different levels of prompting

**How many children were assessed using gradual prompting or task modification procedures?** (Considering the last 40 children who were assessed)

*[Likert scale response]*

**Supplementary Appendix 5.2.** Table of Categorised Assessments

| Assessment name (Standardised administration**): Administered and scored in a consistent manner, which is the same for all students who are assessed. This includes: specific questions/tasks, clear administration guidelines, defined materials and set procedures to elicit responses from the student (assessments do not have to be norm-referenced in order to have standardised administration). *If norm-referenced assessment is not scored according to guidelines then it is not norm-referenced. ** If standardised assessment is not administered to guidelines then it is not standardised | Norm-referenced | Criterion Referenced or Descriptive | Discrete Skill Test (De-contextualised) | Discourse / Text (Contextualised) | Assessment in daily activities | Conducted by SLP | Conducted by other (Not SLP) | Interviews & questionnaires | Clinical Context | School Context | Home/ Community Context |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CASL (Comprehensive Assessment of Spoken Language) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| CCC-2 (Children's Communicative Context – 2nd Edition) | ✓* | X | ✓ | X | X | X | X | ✓ | X | If teacher fills out checklist | If parent fills out checklist |
| CELF:P-2 (Clinical Evaluation of Language Fundamentals: Preschool – 2nd Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| CELF-4 Core Language or Language Index subtests (Clinical Evaluation of Language Fundamentals – 4th Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| CELF-4 Working Memory Index subtests (Clinical Evaluation of Language Fundamentals – 4th Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| CELF-4 Pragmatics Profile (Clinical Evaluation of Language Fundamentals – 4th Ed) | X | ✓ | ✓ | X | X | X | X | ✓ | X | If teacher fills out checklist | If parent fills out checklist |
| CELF-4 Observational Rating Scale (Clinical Evaluation of Language Fundamentals – 4th Ed) | X | ✓ | ✓ | X | X | X | X | ✓ | X | If teacher fills out checklist | If parent fills out checklist |
| CELF-4 Screening Test (Clinical Evaluation of Language Fundamentals – 4th Ed) | X | ✓ | ✓ | X | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| CELF-5 Core Language or Language Index subtests (Clinical Evaluation of Language Fundamentals – 5th Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| CELF-5 Reading Comprehension and Structured Writing (Clinical Evaluation of Language Fundamentals – 5th Ed) | ✓* | X | X | ✓ | X | ✓ | X | X | ✓ | X | X |

| Assessment name (Standardised administration**): Administered and scored in a consistent manner, which is the same for all students who are assessed. This includes: specific questions/tasks, clear administration guidelines, defined materials and set procedures to elicit responses from the student (assessments do not have to be norm-referenced in order to have standardised administration). *If norm-referenced assessment is not scored according to guidelines then it is not norm-referenced. ** If standardised assessment is not administered to guidelines then it is not standardised | Norm-referenced | Criterion Referenced or Descriptive | Discrete Skill Test (De-contextualised) | Discourse / Text (Contextualised) | Assessment in daily activities | Conducted by SLP | Conducted by other (Not SLP) | Interviews & questionnaires | Clinical Context | School Context | Home / Community Context |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CELF-5 Pragmatics Profile (Clinical Evaluation of Language Fundamentals – 5th Ed) | ✓* | X | ✓ | X | X | ✓ | X | ✓ | X | If teacher fills out checklist | If parent fills out checklist |
| CELF-5 Observational Rating Scale (Clinical Evaluation of Language Fundamentals – 5th Ed) | X | ✓ | ✓ | X | X | ✓ | X | ✓ | X | If teacher fills out checklist | If parent fills out checklist |
| CELF-5 Screening Test (Clinical Evaluation of Language Fundamentals – 5th Ed) | ✓* | X | ✓ | X | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| CN-REP (Children's Test of Non-Word Repetition) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| CTOPP-2 (Comprehensive Test of Phonological Processing – 2nd Edition) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| EOWPVT (Expressive One Word Picture Vocabulary Test) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| ERRNI (Expressive, Receptive and Recall Narrative Instrument) | ✓* | X | X | ✓ | X | ✓ | X | X | ✓ | X | X |
| ICPALER Oral Language Supporting Early Literacy Checklists (in Brief or In Depth) checklists | X | ✓ | ✓ | X | X | X | X | ✓ | X | If teacher fills out checklist | If parent fills out checklist |
| NEALE or NARA (Neale Analysis of Reading Assessment) | ✓* | X | X | ✓ | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| NRDLS (Reynell Developmental Language Scales -4th Edition) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| Oral Narrative Assessment Package (by South Australian Government) | X | ✓ | X | ✓ | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| OWLS-II Listening and Speaking tests (Oral and written language scales – 2nd Edition) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |

| Assessment name (Standardised administration**): Administered and scored in a consistent manner, which is the same for all students who are assessed. This includes: specific questions/tasks, clear administration guidelines, defined materials and set procedures to elicit responses from the student (assessments do not have to be norm-referenced in order to have standardised administration). *If norm-referenced assessment is not scored according to guidelines then it is not norm-referenced. ** If standardised assessment is not administered to guidelines then it is not standardised | Norm-referenced | Criterion Referenced or Descriptive | Discrete Skill Test (De-contextualised) | Discourse / Text (Contextualised) | Assessment in daily activities | Conducted by SLP | Conducted by other (Not SLP) | Interviews & questionnaires | Clinical Context | School Context | Home/ Community Context |
|---|---|---|---|---|---|---|---|---|---|---|---|
| OWLS-II Reading and writing Tests (Oral and written language scales – 2nd Edition) | ✓* | X | X | ✓ | X | ✓ | X | X | ✓ | X | X |
| POM (Pragmatics Observational Rating Scale) | ✓* | X | X | X | ✓ | When SLP conducts test | When Other conducts test | X | X | ✓ | ✓ |
| PAT-R 4 (Progressive Achievement Tests in Reading – 4th edition) – when reading comprehension is completed | ✓* | X | X | ✓ | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| Peter & The Cat Retell (by Black Sheep Press) | X | ✓ | X | ✓ | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| PIPA (Pre-Reading Inventory of Phonological Awareness) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| PLS-5 (Preschool Language Scales – 5th Edition) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| PLS-5 Screening Test (Preschool Language Scales – 5th Edition) | ✓* | X | ✓ | X | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| PPVT-4 (Peabody Picture Vocabulary Test – 4th Edition) | ✓* | X | ✓ | X | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| RAPT (Renfrew Action Picture Test) | ✓* | X | ✓ | X | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| RBS (Renfrew Bus Story) | ✓* | X | X | ✓ | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| SALT (Systematic Analysis of Language Transcripts) | X | ✓ | X | ✓ | X | ✓ | X | X | ✓ | X | X |
| SPAT-R (Sutherland Phonological Awareness Test – Revised) | ✓* | X | ✓ | X | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |

| Assessment name (Standardised administration**): Administered and scored in a consistent manner, which is the same for all students who are assessed. This includes: specific questions/tasks, clear administration guidelines, defined materials and set procedures to elicit responses from the student (assessments do not have to be norm-referenced in order to have standardised administration). *If norm-referenced assessment is not scored according to guidelines then it is not norm-referenced. ** If standardised assessment is not administered to guidelines then it is not standardised | Norm-referenced | Criterion Referenced or Descriptive | Discrete Skill Test (De-contextualised) | Discourse / Text (Contextualised) | Assessment in daily activities | Conducted by SLP | Conducted by other (Not SLP) | Interviews & questionnaires | Clinical Context | School Context | Home/ Community Context |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Squirrel story Retell by Black Sheep Press | ✓* | X | X | ✓ | X | ✓ | X | X | ✓ | X | X |
| TAPS-3 (Test of Auditory Perception Skills – 3rd Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| TASP (Test of Aided Language Symbol Performance) | X | ✓ | ✓ | X | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| TILLS (Test of Integrated Language and Literacy Skills) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| TNL (Test of Narrative Language) | ✓* | X | X | ✓ | X | ✓ | X | X | ✓ | X | X |
| TNL-2 (Test of Narrative Language – 2nd Ed) | ✓* | X | X | ✓ | X | ✓ | X | X | ✓ | X | X |
| TOLD-I:4 (Test of Oral Language Development – Intermediate: 4th Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| TOLD-P:4 (Test of Oral Language Development – Primary: 4th Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| TOPL-2 (Test of Pragmatic Language – 2nd Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| TOPS-3 (Test of Problem Solving – 3rd Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| TROG-2 (test of Reception of Grammar – 2nd Ed) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| WJIVOL (Woodcock Johnson IV Ed. Tests of Oral Language) | ✓* | X | ✓ | X | X | ✓ | X | X | ✓ | X | X |
| YARC Early Reading (York Assessment of Reading Comprehension) | ✓* | X | ✓ | X | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |
| YARC Passage Reading (York Assessment of Reading Comprehension) | ✓* | X | X | ✓ | X | When SLP conducts test | When Other conducts test | X | ✓ | X | X |

| Name of assessment (non-standardised administration & scoring): These assessments are not intended to be administered the same way by different assessors in different conditions i.e. procedures for administration & scoring may be variable or may not be described well enough for consistent administration and scoring. | Norm-referenced | Criterion – ref or Descriptive | Discrete Skill Test (De-contextualised) | Discourse/Text (Contextualised) | Assessment in Daily Activities | Conducted by SLP | Conducted by other (Not SLP) | Interviews & questionnaires | Clinical Context | School Context | Home/Community Context |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Non-standardised tests developed by individual SLPs (i.e. short answer test that is highly directed by SLP) | X | ✓ | ✓ | X | X | When conducted by SLP | When conducted by Other | X | If using clinical materials (not context specific) | If using curriculum materials | If using materials from home |
| Standardised tests that have been altered or not administered to guidelines | X | ✓ | ✓ | X | X | When conducted by SLP | When conducted by Other | X | ✓ | X | X |
| Sampling or observing behaviours during role-plays of real-life scenarios | X | ✓ | X | ✓ | X | When conducted by SLP | When conducted by Other | X | If using general scenarios | If using classroom scenarios | If using home scenarios |
| Non-standardised oral language sampling (e.g. narrative, expository, conversation) | X | ✓ | X | ✓ | X | ✓ | X | X | If using clinical materials (not context specific) | If using curriculum materials | If using materials or people from home |
| Non-standardised written language sampling (e.g. reading a passage of text, writing a report) | X | ✓ | X | ✓ | X | ✓ | X | X | If using clinical materials (not context specific) | If using curriculum materials | If using materials or people from home |
| Observations of behaviours/performance in regular class activities | X | ✓ | X | X | ✓ | ✓ | X | X | X | ✓ | X |
| Observations at home or during other regular daily activities (not in class) | X | ✓ | X | X | ✓ | ✓ | X | X | X | X | ✓ |

| Name of assessment (non-standardised administration & scoring): These assessments are not intended to be administered the same way by different assessors in different conditions i.e. procedures for administration & scoring may be variable or may not be described well enough for consistent administration and scoring. | Norm-referenced | Criterion –ref or Descriptive | Discrete Skill Test (De-contextualised) | Discourse/Text (Contextualised) | Assessment in Daily Activities | Conducted by SLP | Conducted by other (Not SLP) | Interviews & questionnaires | Clinical Context | School Context | Home/ Community Context |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Interview or questionnaire with a teacher or school staff member that does not follow a set protocol or have set scoring | X | ✓ | If asking about discrete skills or behaviours | If asking about specific communication functions or contexts | If asking about performance/ participation in daily activities | X | X | ✓ | X | ✓ | X |
| Interview or questionnaire with a parent or caregiver that does not follow a set protocol or have set scoring | X | ✓ | If asking about discrete skills or behaviours | If asking about specific communication functions or contexts | If asking about performance/ participation in daily activities | X | X | ✓ | X | X | ✓ |
| Interview or questionnaire with another person (not caregiver or teacher) that does not follow a set protocol or have set scoring | X | ✓ | If asking about discrete skills or behaviours | If asking about specific communication functions or contexts | If asking about performance/ participation in daily activities | X | X | ✓ | X | X | ✓ |

**Supplementary Appendix 5.3**.

Number of SLPs surveyed in relation to estimated population for each Australian

State/Territory

| State | Number of members per state (data from SPA as at June 2018) | Estimated number of qualified SLPs per state (83% of SLPs are registered with SPA) | Estimated number of SLPs working with children aged 4-12 years (53% of all SLPs) | Number of survey respondents who indicated they work with children aged 4-12 years | SLP participation rate expressed as number of SLPs who participated in the survey against estimated number of SLPs working with children aged 5-12 years (expressed as percentage) |
|---|---|---|---|---|---|
| Australian Capital Territory | 78 | 94 | 50 | 15 | 30.1% |
| New South Wales | 2,160 | 2,602 | 1379 | 135 | 9.8% |
| Northern Territory | 43 | 52 | 27 | 11 | 40.1% |
| Queensland | 1,559 | 1,878 | 996 | 144 | 14.5% |
| South Australia | 586 | 706 | 374 | 46 | 12.3% |
| Tasmania | 108 | 130 | 69 | 14 | 20.3% |
| Victoria | 1,882 | 2,267 | 1202 | 93 | 7.7% |
| Western Australia | 803 | 967 | 513 | 67 | 13.1% |
| **Totals** | **7,219** | **8,698** | **4,610** | **525** | **11.4%** |

**Chapter 6.**

**Language Assessment Practices for Primary School Children (Part II):**

**What Reasons Drive SLPs Choice?**

**Background to Chapter 6 (Journal Article 5)**

Chapter 6 relates to research area three. This chapter provides further information on Australian SLP assessment practice through presentation of data from Part II of the same survey described in Chapter 5. Part II of the survey investigated the language measures, assessment procedures and assessment delivery methods SLPs use and the domains assessed, purposes of use and reasons for which regularly used language measures are chosen. The manuscript contained in this chapter is currently under review with the following journal: Journal of Communication Disorders.

# Language Assessment for Primary School Children (Part II):

# What Reasons Drive SLP Choice?

Deborah Denman[1*], Reinie Cordier[1,4], Natalie Munro[2,1], Jae-Hyun Kim[3,1] and Renée Speyer[4,5,1]

[1] School of Occupational Therapy, Social Work and Speech Pathology, Faculty of Health Sciences, Curtin University, Perth, Australia

[2] Faculty of Health Sciences, The University of Sydney, Sydney, Australia

[3] Department of Linguistics, Macquarie University, Sydney, Australia

[4] Department Special Needs Education, University of Oslo, Oslo, Norway

[5] Department of Otorhinolaryngology and Head and Neck Surgery, Leiden University Medical Centre, Leiden, The Netherlands

*Corresponding author: E-mail: deborah.denman@postgrad.curtin.edu.au

Keywords: language disorder, speech pathology, assessment, survey, children

## 6.1. Abstract

Background: Intervention services for primary school children with language disorder are informed by assessment findings, therefore it is important that the language measures and assessment procedures that SLPs choose to use are effective in identifying the needs of children.

Objective: This study reports on data from Part II of a survey of SLP language assessment practices for school-aged children. The objective of the study was to identify the specific language measures, assessment procedures and assessment delivery methods used by SLPs to assess the language abilities of primary school children and the domains targeted, purposes of use and reasons for which language measures were chosen for use.

Methods: A total of 335 SLPs provided information in a web-based survey regarding the regularity with which they use different assessment measures, procedures and methods. For regularly used language measures, SLPs were also asked to identify the domains targeted, purposes of use and reasons for which measures were chosen for use. Terms and definitions from a recently developed taxonomy were used to guide the development of survey questions and provide explicit terminology for describing the domains and purposes assessed.

Results: SLPs collectively listed a large array of language measures and assessment procedures, although only a small number of measures were used regularly by each SLP. One in three SLPs reported that other personnel conduct assessments on SLPs' behalf and one in sixteen reported using information and communication technologies as methods of conducting assessment. SLPs favoured language measures that target semantics and syntax in word and sentence level tasks over measures that target social abilities and discourse. SLPs appeared to select diagnostic measures based on psychometric properties, but not screening measures.

Conclusion: Findings from this study indicate that SLPs need to give greater consideration to evidence-based practice recommendations when assessing the language abilities of primary school children. Implications for clinical practice and future directions are discussed.

**6.2. Introduction**

Children with language disorder present with ongoing difficulties learning spoken and written language compared to peers, with these difficulties having a significant impact of daily functioning (Bishop, Snowling, Thompson, Greenhalgh, & CATALISE-2 consortium, 2017). Approximately 10% of primary school children have a language disorder, placing them at high risk for social-emotional problems, behavioural difficulties, and poor progress with literacy and numeracy (Harrison, McLeod, Berthelsen, & Walker, 2009; Norbury et al., 2016). To identify appropriate interventions and educational supports for these children, SLPs must first assess a child's language abilities. Given the importance of language assessment data in determining service provision, it is important that SLPs make evidence-based decisions when assessing the language abilities of children (Betz, Eickhoff, & Sullivan, 2013). Despite this, limited information exists regarding the specific language measures, assessment procedures and assessment delivery methods SLPs use with primary school children and the reasons for which these are chosen. Information on current SLP assessment practice is needed to better understand the alignment between current practice and evidence-based practice which, in turn, may assist in identifying future actions to improve clinical practice in the field of child language assessment (Eadie, 2003).

**6.2.1. Language measures and assessment procedures**. Language measures have set guidelines for consistent administration, including specific questions or tasks, defined scoring guidelines, and set procedures to elicit responses from the child (Hegde & Pomaville, 2017). These language measures allow for a child's performance to be compared to other children, since all children undertake the assessment tasks under the same conditions (Betz et al., 2013). In contrast, assessment procedures do not have specific guidelines for consistent administration, but instead allow for data to be collected on a child's performance in natural everyday environments (Hegde & Pomaville, 2017). Due to the different information

supplied by each, it is recommended that SLPs use a combination of language measures with set guidelines for administration and scoring and assessment procedures without set guidelines for administration and scoring when assessing the language abilities of primary school children (Bishop, Snowling, Thompson, & Greenhalgh, 2016).

Language assessment may also be conducted using different service methods. Traditionally, assessment is conducted through direct testing, sampling or observation by the SLP; however, assessment could also be conducted through testing, sampling or observation by other personnel such as other health professionals, teachers or therapy assistants (Denman, Kim, Munro, Speyer, & Cordier, 2019). Although some literature exists to support delivery of language intervention by other personnel who have been trained (Boyle, McCartney, Forbes, & O'Hare, 2007), no studies have explicitly investigated the practice of having other personnel conduct assessments on the SLPs behalf. Therefore, no guidelines currently exist to support this practice.

Language assessment may also be conducted via information and communication technologies (ICTs) or proxy-report (Denman et al., 2019). The use of ICTs has been shown to be effective for delivering language assessment and this method has significant potential to increase access to assessment services and reduce travel time for children and families (Mashima & Doarn, 2009; Waite, Theodoros, Russell, & Cahill, 2010). Assessment data collected via proxy-report are typically interviews or questionnaires (Dockrell & Marshall, 2015). Proxy-reported methods provide an avenue for collecting data on a child's functional performance with everyday communication partners and also allow for families to be included in the assessment process (Bishop et al., 2016; Roulstone, 2015). Data from proxy-reported methods is particularly important for the development of intervention goals that are aligned with functional needs (Bishop et al., 2016; Trembath, Westerveld, & Shellshear, 2016).

**6.2.2. Language assessment domains and purposes**. When assessing primary school children with language disorder, SLPs may target a number of different domains, including semantics, morphosyntax, social abilities and discourse, meta-abilities and executive functions (Denman et al., 2019; Dockrell & Marshall, 2015). Measures that target social abilities and discourse should be conducted alongside measures that target other domains, as these assessments often give different but valuable information on a child's abilities (Ebert & Scott, 2014; Lennox, Westerveld, & Trembath, 2018). Furthermore, domains should be assessed using measures that are well-suited for measuring the specific domains of interest. For example, social abilities should be assessed in situations that are reflective of naturalistic, everyday communication environments (Volden et al., 2017).

Assessment may also be conducted for a variety of purposes including predicting outcome, selecting intervention, planning dosage, diagnosis, screening, detecting change or describing status (Denman et al., 2019; Dockrell & Marshall, 2015; Wade, 2004). It is important that language measures are well-matched to the purposes for which assessment data are to be used (Wade, 2004). For example, measures used for screening and diagnostic purposes should have sound psychometric properties, including diagnostic accuracy; measures used for detecting change should have evidence of being responsive to change; and measures used for selecting intervention goals should target performance in real life situations (Denman et al., 2017; Ebert & Scott, 2014; Polit, 2015). Collectively, using different language measures and assessment procedures that target appropriate domains and using measures for the purposes they have been developed; all contributes towards formulating an accurate understanding of a child's language abilities.

**6.2.3. Current knowledge of SLP assessment practice.** A wide array of measures, procedures and methods exist for assessing the language abilities of primary school children, with one recent systematic review identifying over 70 oral language measures for children

aged 4-12 years (Denman et al., 2017). Despite this proliferation of assessment options, few previous studies have examined the different language measures and assessment procedures that SLPs use for the broad population of primary school children (Beck, 1995; Betz et al., 2013; Caesar & Kohler, 2009; Roulstone et al., 2015; K. S. Wilson, Blackmon, Hall, & Elcholtz, 1991). These studies identified that both language measures with set guidelines for administration and scoring and  assessment procedures without set guidelines for administration and scoring are used by SLPs, although greater focus appears to be placed on use of  language measures with set guidelines for administration and scoring (Caesar & Kohler, 2009; Fulcher-Rood, Castilla-Earls, & Higginbotham, 2018). The most commonly used language measures are comprehensive language measures or single word vocabulary measures (Betz et al., 2013; Caesar & Kohler, 2009). SLPs appear to favour recently published measures that have long histories in the field through multiple editions, such as versions of the Clinical Evaluation of Language Fundamentals (Wiig, Secord, & Semel, 2004; Wiig, Semel, & Secord, 2004, 2013). SLPs also appear to favour measures with locally developed normative data, for example, measures developed in the United Kingdom such as the Renfrew Action Picture Test (RAPT) (Renfrew, 2010) and editions of the Reynell Developmental Language Scales (Letts, Edward, Schaefer, & Sinka, 2014) are commonly used by SLPs in the United Kingdom, but not in the United States of America (Betz et al., 2013; Roulstone et al., 2015).

The most commonly used assessment procedures appear to be parent interviews and assessments described as 'language sampling' or 'observations' (Caesar & Kohler, 2009; Fulcher-Rood et al., 2018; Roulstone et al., 2015; Singh, Chan, & Rusli, 2016). However, the regularity with which SLPs are reported to use these types of assessment procedures is varied across studies. For example some studies have identified that SLPs routinely conduct 'language sampling' (Caesar & Kohler, 2009; Westerveld & Claessen, 2014), while another

study identified that SLPs do not regularly use this assessment procedure (Pavelko, Owens, Ireland, & Hahs-Vaughn, 2016) These differences may be due to differences between survey samples, differences in how SLPs interpret the term 'language sampling', and use of different scales for rating 'regularity of use' (Pavelko et al., 2016).

Although these earlier studies have provided important information regarding SLP assessment practice, notable limitations exist. Firstly, previous surveys were largely conducted in the United States of America (Beck, 1995; Betz et al., 2013; Caesar & Kohler, 2009; Fulcher-Rood et al., 2018; Huang, Hopkins, & Nippold, 1997; K. S. Wilson et al., 1991). Two studies from outside the United States of America have been conducted in the United Kingdom and Malaysia, however, both these studies have small sample sizes which limit generalizability of findings (Roulstone et al., 2015; Singh et al., 2016). Similarly, although two previous Australian surveys of language assessment practice have been conducted, these surveys examined specific aspects of practice (i.e. assessments described as 'language sampling' or assessments specifically for children with culturally or linguistically diverse backgrounds; Westerveld & Claessen, 2014; Williams & McLeod, 2012). No previous studies have been conducted examining Australian SLP's use of both language measures with guidelines for administration and scoring and assessment procedures without set guidelines for administration and scoring when assessing the broader population of primary school children. Therefore, a survey of Australian SLP assessment practice will add to current knowledge in the field and provide information that may assist in identifying trends across countries (Singh et al., 2016).

Secondly, no published studies have examined the regularity with which language assessments for school-aged children are conducted by other personnel. Similarly, although one previously published research study examined the frequency with which SLPs use ICTs when conducting language assessment (Tucker, 2012), this study was limited in scope to

SLPs employed in schools in the United States of America. Identifying the extent with which these assessment delivery methods are used by SLPs from different agencies will assist in further profiling assessment practice and identifying future research directions.

Thirdly, no previous studies have explicitly examined the domains that SLPs target when using specific language measures. Similarly, while previous studies have examined the language measures SLPs selected for diagnostic purposes (Betz et al., 2013; Fulcher-Rood et al., 2018; K. S. Wilson et al., 1991); no previous studies have examined SLP use of language measures for other purposes. This information is important for understanding how language measures are being used by SLPs and to identify if actual use aligns with recommended purposes of use.

Lastly, limited information is available regarding the reasons for which SLPs use language measures. Previous  studies have identified that assessment choice may be driven by service policy, that SLPs may not be selecting language measures based on psychometric quality, and that time constraints related to large caseloads may lead to the selection of language measures that are quick to administer and allow for multiple domains to be assessed simultaneously (Beck, 1995; Fulcher-Rood et al., 2018; Huang et al., 1997; K. S. Wilson et al., 1991). However, as these studies have focused on SLPs who work in school-based service agencies in the United States of America, it is not known if the findings apply to SLPs who work in different settings. Only one study examined SLPs from diverse work agencies, however, this study was confined to one state in the United States of America and was published more than 20 years ago (Huang et al., 1997). Therefore, to better understand current SLP decision-making in child language assessment, further survey research is needed using a broad population of SLPs from different agencies and geographical locations (Fulcher-Rood et al., 2018).

**6.2.4. Objectives**. The objective of this study was to investigate the assessment practices used by Australian SLPs to assess the language abilities of primary school children. Specifically, this study addresses the following research questions:

1. What specific language measures, assessment procedures and assessment delivery methods do SLPs regularly use to assess the language abilities of children aged 4-12 years?

2. For language measures that SLPs regularly use to assess children aged 4-12 years: what are the domains targeted, the purposes of use, and the main reasons why measures are chosen for use?

## 6.3. Method

Ethical approval to conduct this survey was provided by the [deleted for review]. The online survey was created using Qualtrics software (Qualtrics, 2005) and research literature on survey design was used to guide the structure and format of the survey questions (Andrews, Nonnecke, & Preece, 2003).

**6.3.1. Survey structure and format.** The survey consisted of four sections. Section one of the survey was designed for all SLPs in Australia to complete and consisted of multiple-choice questions about participant demographics including: membership with the Australian speech pathology association (Speech Pathology Australia), gender, years since graduation, qualifications, postcode of workplace, current employment status, and client population. The remaining survey sections were for SLPs who indicated in section one that they had provided services in the last year to at least 40 children aged 4-12 years with language disorder. These sections of the survey asked questions about work agency and work context (section two), the regularity with which different assessment delivery methods were used (section three) and the regularity with which different language measures and assessment procedures were used (section four). Section four also asked about the domains assessed with regularly used language measures, the purposes for which regularly used

language measures were used and the reasons for which regularly used language measures were selected. A copy of the survey questions related to this study (section one to four) is provided in Supplementary Appendix 6.1. The survey also included questions relating to the regularity with which SLPs collect different types of data, assess in different tasks and contexts and use dynamic assessment procedures, however, findings related to these questions are reported in an accompanying publication (Denman, Cordier, Kim, Munro, & Speyer, Under Review).

Assessment procedures and assessment delivery methods were listed in the survey and participants were asked to indicate the regularity with which they use each procedure or method using a 5-point Likert Scale. For language measures, participants were asked to write the measures that they use by name in open text boxes and select the regularity with which they use each measure on the same 5-point Likert Scale. To facilitate consistent application of the frequency rating scale by survey participants, the Likert scale points were associated with numeric qualifiers, as well as descriptive terms (Blais & Grondin, 2011). Participants were asked: "How many children were assessed using the [measure, procedure or method] considering the last 40 children who were assessed?" The response options included 'no children', 'few children' (i.e. 5 or less), 'some children' (i.e. 6-19), 'many children' (i.e. 20-34) or 'most children' (i.e. 35 or more). For language measures that participants indicated using regularly, participants were then asked to indicate up to three domains that were primarily assessed when the measure was used, up to three main purposes for which the measure was used and up to three main reasons why the measure was chosen for use.

To ensure SLPs applied the same definitions when describing delivery methods, domains and purposes; terms and definitions from a recently developed taxonomy were used in the survey questions (Denman et al., 2019). This taxonomy provided explicit terminology for describing language assessment practices with this terminology previously agreed upon

by over 40 Australian SLPs experienced in the field of child language using a Delphi consensus method. Specifically, the survey questions in this study relate to assessment domain (Aspect I of the taxonomy), assessment purpose (Aspect II of the taxonomy), assessment method (a component of aspect III of the taxonomy) and presence of standardisation (a component of Aspect IV of the taxonomy). The taxonomy terms and definitions that were used in this survey are supplied in Table 6.1. During completion of the survey, SLPs were instructed, that although terms may be used differently, they must use definitions provided in the survey when answering the questions.

Prior to dissemination of the survey, SLPs from four different agencies trialed the survey and provided feedback. Survey completion time was estimated as five minutes for SLPs completing only section one and between 25-40 minutes for SLPs completing all four sections. Skip logic was used throughout the survey so that participants were only presented with questions that were relevant to them based on their previous answers. Participants were also able to complete the survey in more than one sitting as survey responses could be saved and re-opened later.

The survey was accessible between mid-February and mid-June 2018. The link to the survey was distributed through Twitter, Face book posts and via the national Speech Pathology Australia newsletter distributed to all association members. The survey link was also emailed to numerous SLPs through publicly available email addresses, email discussion groups, and the professional networks of the researchers. SLPs who received the link were encouraged to disseminate around their professional networks.

Table 6.1

*Taxonomy terms (with definitions and examples) for describing assessment methods, domains*

*and purposes (from Denman et al., 2019)*

---

**Assessment Domains (Taxonomy Aspect I)**

---

**Semantics**

Words and word meanings i.e. vocabulary, word retrieval, lexical meaning, word definitions.

Examples:

- *Vocabulary knowledge is assessed using a picture-naming task*
- *A sample of a child's language is analysed for or type-token ratio (TTR) or number of different words spoken (NWD)*

---

**Morphosyntax**

Different word forms and the order and combination of words in sentences i.e. syntax and morphology.

Examples:

- *Knowledge of sentence structure is assessed by asking a child to point to pictures that represent a spoken sentence*
- *A sample of a child's language is analysed for Mean Length of Utterance (MLU) or Brown's Grammatical Morphemes*

---

**Social Abilities and Discourse**

Giving and making meaning in social context or communication for social purposes. Includes:

- Pre-linguistic communication e.g. joint attention, gesturing
- Communication intentions/purposes e.g. requesting, commenting, greetings, asking questions, giving reasons
- Non-verbal communication e.g. body language and facial expressions
- Non-literal language e.g. jokes, sarcasm, metaphors, inferences
- Matching communication style to social context e.g. adjusting communication style between friends and teachers
- Conversation conventions e.g. topic selection and maintenance, conversational turn-taking
- Text cohesion e.g. presence of mazes or incomplete sentences
- Text organisation e.g. narrative structure (story grammar), episodic structure etc.

Examples:

- *Text cohesion and organisation are assessed during a narrative retell task*
- *A teacher completes a checklist of the conversational conventions that a child demonstrates in day-to-day conversations with teachers and peers*

---

**Meta-Abilities**

Ability to reflect on own thinking processes and learning and understand how to regulate these processes. Includes:

- Meta-cognition: Knowledge and use of strategies for self-monitoring and managing own learning
- Meta-language: Knowledge of phonemic (phonemic awareness), morphological/syntactic (meta-syntactic) or text-level (meta-narrative) rules and an ability to effectively apply these rules for improved performance
- Meta-pragmatics: Knowledge of social conventions in relation to own communication and ability to apply this knowledge for improved performance

Examples:

- *Meta-narrative abilities are assessed by asking a child to describe their understanding of what constitutes good narrative structure*
- *Phonemic awareness (meta-language) abilities are assessed by asking the child to identify the sounds they hear in words*

---

**Executive Functions**

Cognitive processes required for execution of purposeful, controlled and goal-oriented behaviour. Includes:
- Inhibition: Ability suppress inappropriate thoughts, comments and behaviours in order to focus and attend to tasks i.e. self-control
- Emotion control: Ability to manage emotions in order to achieve goals or and complete tasks i.e. self-regulation
- Working memory: Ability to retain, process and manipulate pieces of information for short periods of time in order to complete required tasks
- Organisation: Ability to use organisational strategies for task completion i.e. strategic planning
- Mental flexibility: Ability to integrate prior knowledge and experiences or effectively apply of different rules for different situations
- Sustained attention: Ability to maintain attention despite distractions and fatigue in order to complete tasks

Examples:
- *Working memory is assessed through a series of number repetition tasks*
- *Organizational abilities are assessed by observing a child plan out a project and identify the steps involved in completing the project*

## Assessment Purposes (Taxonomy Aspect II)

**Predict outcome**

Identify risk, predict need for intervention or identify support strategies

Examples:
- *Curriculum differentiation or education supports are identified by assessing the child's performance in the presence of different prompts or scaffolds (e.g. visual supports versus no visual supports)*
- *Preschool children are assessed on pre-literacy skills that may be predictive of later literacy success, with results used to identify those who may benefit from participation in a preventive program*

**Select intervention**

Identify suitable intervention approaches or select intervention goals/targets.

Examples:
- *A parent interview assists with selection an intervention approach by identifying family preferences, child's likes/dislikes and available resources*
- *A child's ability to produce sentences of varying complexity is assessed to identify intervention targets*

**Plan dosage**

Predict intensity (dosage) of intervention.

Examples:
- *Undertaking a short trial of the intervention to assess the child's response to intervention*
- *Interviewing school staff regarding previous interventions that have been implemented and the child's response to these*

**Screening**

Identify children who may have a disorder that requires further diagnostic assessment to confirm.

Examples:
- *A selection of skills are screened to determine if a child should undergo further diagnostic assessment and the domains that should be targeted in further assessment*

**Diagnostic**

Diagnose a condition by making a comparison with peers.

Examples:
- *A child is assessed to determine the presence or severity of language disorder; or determine if functioning is different to peers*

**Detect change**

Measure change in status or monitor progress over time.

Examples:

- *Recording a 'baseline' for future comparison*
- *Repeating the same assessment at different intervals to monitor progress*

**Describe status**

Describing or profiling a particular aspect of a child's functioning.

Examples:

- *A profile of communicative behaviours (gesture dictionary) is created to describe communication behaviours*
- *Narrative retell skills are described in terms of strengths and weaknesses*

## Assessment Methods (Component of Taxonomy Aspect III)

**By Person - Conducted by SLP**

Assessment conducted by an SLP through pre-planned observation, testing or sampling of a child's skills. Results may be analysed at the time or may be analysed later from an audio/video recording. Other people may assist with administration; however, the SLP has the primary role in planning the assessment and analysing findings

Examples:

- *An SLP conducts a standardised assessment with a child*
- *An SLP transcribes and analyses a narrative sample that was audio-recorded earlier by a teacher in class*

**By Person - Conducted by Other**

Assessment conducted by another person (teacher, another professional etc.), through pre-planned observation, testing or sampling of the child's skills. An SLP may provide training or support to the other person however the other person has a key role in planning the assessment and analysing/interpreting results. Note: SLP students who are conducting assessments under supervision of a qualified SLP are not considered to be "other personnel".

Examples:

- *A teacher conducts a standardised assessment with a child*
- *A teacher observes a child in the playground and makes notes on behaviours that are observed*

**Proxy-Report**

Skills are documented based on retrospective reports from others (as opposed to being documented in the moment they occur). The reported information: may be from the child (self-report), a caregiver, a teacher or a peer and information may relate to previous skills or current abilities.

Examples:

- *An SLP undertakes a parent interview to collect information from a parent on aspects of a child's history that may be diagnostically significant*
- *A teacher completes a checklist regarding the behaviours they have observed the child use during the previous school term*

**Face-to-face (only applies to assessments conducted by a person)**

Assessment is conducted with the child and an assessor in the same room

Examples:

- *During a face-to-face interaction with a child, an SLP audio-records a language sample for later analysis*
- *An SLP or another professional administer a test face-to-face*

**ICT (only applies to assessments conducted by a person)**

Assessment is conducted with the assessor and the child communicating through ICTs (Information and Communication Technologies). Technology that is not used for two-way communication between individuals during the assessment (e.g. audio/video recorders or scoring software) is not considered ICT

Examples:

- *An SLP or another professional administer a test using Skype, Zoom, video-conferencing, telephone or other form of communication technology*

---

**Presence of Standardisation (Component of Taxonomy Aspect IV)**

---

**Standardised**

Assessments intended to be conducted and scored in a consistent manner i.e. set questions or tasks, specific administration and scoring instructions, prescribed assessment materials and specific procedures to elicit responses from the child.

- *An SLP conducts an assessment that follows specific administration and scoring guidelines*

---

**Non-standardised**

Assessments that may not be conducted the same way by different assessors. Procedures for administration and scoring may be variable or may not be described sufficiently to ensure that assessment is conducted in the same way each time.

- *An SLP conducts an assessment that they designed themselves to collect specific data needed on a particular child's language abilities*

---

**6.3.2. Data analysis.** Data from the survey was imported from Qualtrics into the Statistical Package for the Social Sciences (SPSS) version 20 program (IBM Corp, Released 2011). Descriptive statistics were used to report on the regularity with which different language measures, assessment methods and service methods were used by SLPs and the domains, purposes and reasons for which SLPs reported using standardised measures. For the purposes of survey analysis, 'regular use' was defined as being used for 'many' or 'most' children (i.e., with half or more than half of the last 40 children who received services).

As SLPs typically only used one edition of a language measure and did not always specify the edition used, different editions of the same language measure were counted as one standardised measure during analysis. The exception was versions of the Clinical Evaluation of Language Fundamentals (CELF) and Test of Narrative Language (TNL). As some SLPs identified using two different editions of these measures, different editions of these particular measures were counted as two different measures. Any responses listed by SLPs that were not measures of language, for example speech production measures, were removed from analysis. A total of 50 (2.2%) of responses were also removed from the analysis as they could not be identified from information supplied in the survey. This included responses that referred to a general procedure, such as 'language sampling' or 'language screener' or acronyms that were ambiguous or could not be identified in online Google searches.

**6.4. Results**

      **6.4.1. Survey participants.** In total, 847 survey responses were received, with 727 being complete and valid survey responses (85.8% total completion rate). Of the completed responses, 83.4% of SLPs identified themselves as being current members of the national speech pathology association, Speech Pathology Australia. This estimate is comparable to the 80% membership estimate obtained in a previous Australian survey of SLPs (Westerveld & Claessen, 2014).

      Of the SLPs who completed the survey, 525 SLPs identified themselves as working with children 4-12 years with language disorder. From Speech Pathology Australia membership data, it was estimated that approximately 53% of Speech Pathology Australia members who are qualified SLPs work with primary school children (L. Young, personal communication, 4th June and 20th September, 2018). Using 83% as an estimate of association membership and 53% as an estimate of the proportion of SLPs who work with primary school children, the number of SLPs in Australia who work with primary school children was estimated as 4,610 at the time of the survey. Therefore, the 525 responses in this survey represent approximately 11.4% of the estimated target population size. Further details on the sample population size are provided in the publication that accompanies this publication (Denman et al., Under Review).

      **6.4.2. Participant demographics.** Of the 525 SLPs who indicated in the survey that they worked with primary school children, 407 SLPs indicated having provided a service to 40 or more children with language disorder in the preceding year. Of this 407, 335 SLPs completed all the survey questions (82.3% completion rate for SLPs who indicated frequently providing services to primary school children with language disorder). The data provided by these 335 SLPs is reported on in this study. Demographic data reported by these SLPs is outlined in Table 6.2. No significant differences were identified between the 335 SLPs who

completed the survey and the 72 SLPs who indicated having serviced at least 40 children in the last year with language disorder but did not complete the survey with regards to service agency $X^2$ (5, N=407) = 9.055, p =0.107, Australian state/territory $X^2$ (7, N=407) = 11.13, p =0.133 or years since graduation $X^2$ (4, N=407) = 5.86, p =0.210.

Table 6.2

*Demographics of Survey Participants Who Work with Children Aged 4-12 Years with Language Disorder and Completed the Survey (Survey Part II; n=335)*

| Category | Subcategory | Total (%) |
| --- | --- | --- |
| Gender | Female | 323 (96.4) |
| | Male | 11 (3.3) |
| | Other | 1 (0.3) |
| | **Total** | **335 (100)** |
| Australian State | ACT | 7 (2.1) |
| | NSW | 79 (23.6) |
| | NT | 7 (2.1) |
| | QLD | 101 (30.1) |
| | SA | 31 (9.3) |
| | TAS | 13 (3.9) |
| | VIC | 55 (16.4) |
| | WA | 42 (12.5) |
| | **Total** | **335 (100)** |

| Category | Subcategory | Total (%) |
|---|---|---|
| Agency through which service is provided | [a] Education agency | 117 (34.9) |
| | [b] Private practice | 104 (31.0) |
| | [c] Health agency | 50 (14.9) |
| | [d] Disability specific agency | 43 (12.8) |
| | [e] General agency | 15 (4.5) |
| | [f] University | 6 (1.8) |
| | **Total** | **335 (100)** |
| [g] Remoteness of geographical location | Regional/Remote | 109 (32.5) |
| | Major City (Metropolitan) | 226 (67.5) |
| | **Total** | **335 (100)** |
| Years since graduation | 21+ years | 71 (21.2) |
| | 11-20 years | 68 (20.3) |
| | 6-10 years | 66 (19.7) |
| | 3-5 years | 84 (25.1) |
| | 0-2 years | 46 (13.7) |
| | **Total** | **335 (100)** |
| Frequency of children on caseload from Culturally and Linguistically Diverse (CALD) backgrounds | [h] High frequency of CALD | 52 (15.5) |
| | [i] Low frequency of CALD | 283 (84.5) |
| | **Total** | **335 (100)** |

Note: [a] Education agency i.e. education department or school (may be government or non-government); [b] Private practice i.e. business owner or employee in private practice; [c] Health agency i.e. health department or hospital (may be government or non-government); [d] Disability specific agency i.e. children must have diagnosis or suspected diagnosis of disability to access the service (may be government or non-government); [e] General agency i.e. agency that is not identified as other category (may be government or non-government); [f] University i.e. student teaching clinic; [g] As classified by the Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS)*. 2016 [cited 2018 March]; Available from: http://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.005; [h] More than half of children with Language Disorder on caseload were from CALD (Culturally and Linguistically Diverse) backgrounds e.g. bilingualism or standard Australian English is not first language; [i] Less than half of children with Language Disorder on caseload were from CALD (Culturally and Linguistically Diverse) backgrounds e.g. bilingualism or standard Australian as not first language).

**6.4.3. Language measures and assessment procedures**. The 335 SLPs who completed the survey collectively listed 139 identifiable language measures (i.e., measures with set guidelines for administration and scoring) as being used to assess the last 40 children. On average, each SLP listed 6.9 (SD=3.23) different language measures. Six participants indicated not using any language measures with the last 40 children they assessed and two SLPs identified using 15 or more different language measures.

Of the 139 language measures used by SLPs, 69 were identified as being used regularly (i.e. used to assess half or more than half of the last 40 children) by at least one SLP. Only five of the 69 measures were identified as being used regularly by more than 2.4% (8/335) of SLPs. These five measures were the Clinical Evaluation of Language Fundamentals – 4th Edition (CELF-4) core or language index subtests (Wiig, Semel, et al., 2004) (used regularly by 37.3% or 125/335 SLPs), the Clinical Evaluation of Language Fundamentals – 5th Edition (CELF-5) core or language index subtests (Wiig et al., 2013) (used regularly by 17.9% or 60/335 SLPs), the Clinical Evaluation of Language Fundamentals: Preschool – any edition (CELF:P) core language subtests (Wiig, Secord, et al., 2004) (used regularly by 13.7% or 46/335 SLPs), Renfrew Action Picture Test – any edition (RAPT) (Renfrew, 2010) (used regularly by 27.8% or 93/335 SLPs), and the Sutherland Phonological Awareness Test – any edition (SPAT) (Neilsen, 2003) (used regularly by 12.5% or 42/335 SLPs).

The most regularly used assessment procedures (i.e., procedures without set guidelines of administration and scoring) were reported to be interview with parent (used regularly by 36.7% or 123/335 SLPs) and interview with teacher (used regularly by 26.7% or 89/335 SLPs). Approximately one fifth of SLPs (19.7% or 66/335 SLPs) reported not using parent interviews for any children and one quarter (23.0% or 77/335 SLPs) reported not interviewing a teacher for any children (considering the last 40 children assessed).

Observations of abilities in the classroom were reported as being used regularly by 26.9% (or 90/335 SLPs) and observations at home was reported as being used regularly by 11.6% (or 39/335 SLPs). One fifth of SLPs (21.2% or 73/335 SLPs) reported not conducting any classroom observations and over one half (57.3% or 192/335 SLPs) reported not conducting any observations in home or community contexts (considering the last 40 children assessed).

Oral language sampling procedures were reported as being conducted regularly by 15.5% (or 52/335 SLPs). Only 6.3% of SLPs (21/335) reported regularly sampling language in role play and only 4.7% (16/335 SLPs) reported regularly using written language sampling procedures. Two fifths of SLPs (39.1% or 131/335 SLPs) reported not using oral language sampling procedures, over half (58.5% or 96/335 SLPs) reported not sampling language in role play and two thirds 64.8% or 217/335 SLPs) reported not using written language sampling procedures.

The percentage of SLPs who regularly used each language measure (for the five measures used most regularly) and each assessment procedure is shown in Figure 6.1. A further list of the language measures used by SLPs and associated frequencies of use is contained in Supplementary Appendix 6.2.

**6.4.4. Service methods.** Approximately one third of SLPs (29.8% or 100/335) reported that other personnel conduct assessment on their behalf, with 6.6% (or 22/335 SLPs) reporting that other personnel conducted assessment regularly (i.e. with half or more than half of the last 40 children assessed). Of the 100 SLPs who reported other personnel conducting assessment services, 70.0% (or 70/100 SLPs) indicated these other personnel to be teachers, teacher-aides or therapy assistants; 27.0% (or 27/100 SLPs) indicated the other personnel to be other allied health professionals, such as psychologists or occupational therapists; and 3.0% (or 3/100 SLPs) indicated that both of these groups of people conducted assessments.

With regards to ICT, only 6.3% (or 21/335 SLPs) reported using ICTs to deliver assessment services, with only one SLP reporting regular use of ICTs when assessing the language abilities of primary school children.
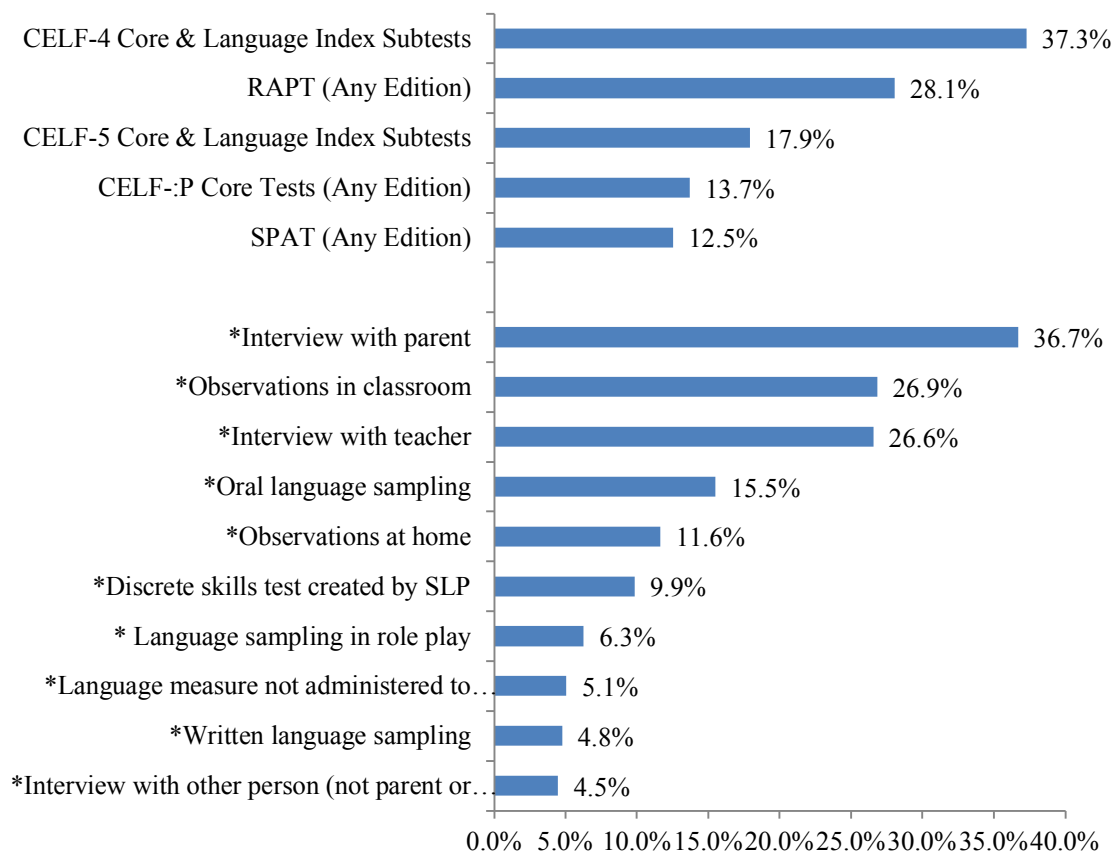


*Figure 6. 1.* Number of SLPs who regularly use each assessment (n=335).

Percentage of SLPs who regularly used each language measure and assessment procedure (n=335). Regular use of a language measure or assessment procedure was defined if the measure or procedure was used with half or more than half of the last 40 children. Language measures and assessment procedures that were used regularly by 2.4% (8/335) or less participants are not included in this figure. *Assessment procedures with no set guidelines for administration or scoring.

**6.4.5. Domains, purposes and reasons for use of language measures**. For language measures that were used regularly (i.e. used with half or more than half of the last 40 children), SLPs selected the main domains (up to three) that were targeted in the language measure, the main purposes for which the measure was used (up to three), and reasons why

the measure was chosen for use (up to three rank ordered). The taxonomy definitions were used as definitions for domains and purposes (see Table 6.1).

With regards to domains, most SLPs reported using the Clinical Evaluation of Language Fundamentals (CELF) core language or language index subtests or the Renfrew Action Picture Test (RAPT) to target semantics and morphosyntax. However, one quarter to one third of SLPs also selected a focus on executive functions when they used versions of the Clinical Evaluation of Language Fundamentals (CELF) core language or language index subtests and one quarter indicated focusing on social-abilities when they used the Clinical Evaluation of Language Fundamentals-4th Edition (CELF-4) core or language index subtests or the Renfrew Action Picture Test (RAPT). The Sutherland Phonological Awareness Test (SPAT) was primarily used to assess meta-abilities, although approximately one fifth of SLPs also selected semantics and morphosyntax as areas they primarily focus on when using this measure. The domains targeted by the five most regularly used language  measures are displayed in Figure 6.2.
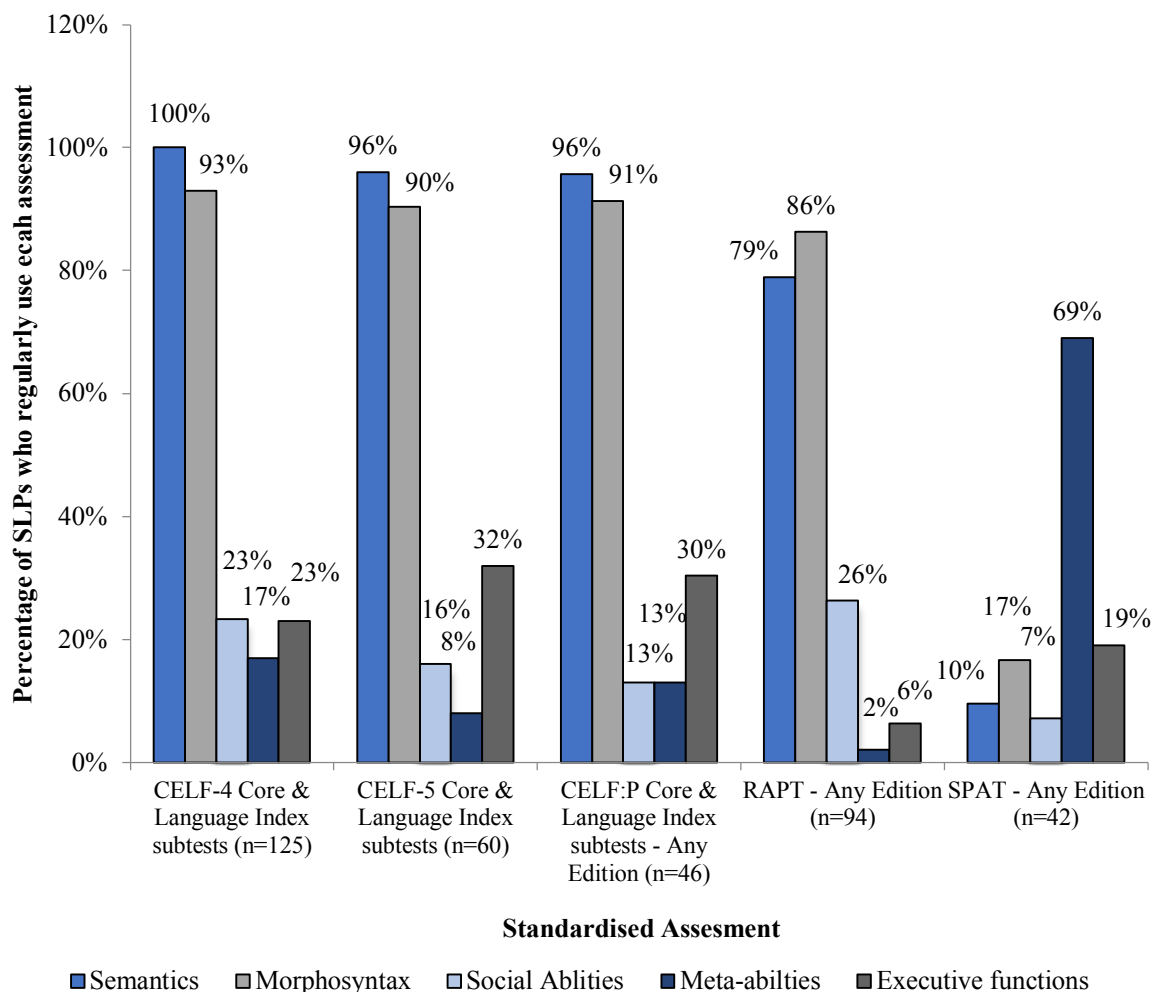
*Figure 6. 2.* Domains targeted by the five most regularly used language measures (n=335).

Language measures have set guidelines for administration and scoring. Regular use of a measure was defined if the measure was used with half or more than half of the last 40 children. SLPs could select up to three main domains they target for each measure. On average, SLPs selected 2.2 domains for each measure.

With regards to purposes, over 80% of SLPs who regularly used versions of the Clinical Evaluation of Language Fundamentals (CELF) core language or language index subtests reported using these measures for diagnostic purposes and approximately half indicated using these measures for predicting outcome, selecting intervention, and describing status. Over 80% of SLPs who regularly used the Renfrew Action Picture Test (RAPT) reported using this measure for screening purposes and approximately half indicated using this measure for the purposes of detecting change. The Sutherland Phonological Awareness

Test (SPAT) was reported as predominantly used for predicting outcome and selecting intervention. The purposes for which regularly used language measures were used can be found in Table 6.3.

Table 6.3

*Assessment Purposes for Regularly Used Language Measures*

| Purpose | CELF-4 Core or Language Index (n=125) | | CELF-5 Core or Language Index (n=60) | | CELF:P Core Subtests (n=46) | | RAPT – Any Edition (n=94) | | SPAT – Any Edition (n=42) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Top reason | Not top reason | Top reason | Not top reason | Top reason | Not top reason | Top reason | Not top reason | Top reason | Not top Reason |
| Predict Outcome | 55.2% | 44.8% | 61.7% | 38.3% | 63.0% | 37.0 | 26.9% | 73.1% | 59.5% | 40.5% |
| Select Intervention | 47.2% | 52.8% | 48.3% | 51.7% | 58.7% | 41.3% | 31.2% | 68.8% | 64.3% | 35.7% |
| Plan Dosage | 4.0% | 96.0% | 1.7% | 98.3% | 4.3% | 95.7% | 1.1% | 98.9% | 2.4% | 97.6 |
| Diagnosis | 83.2% | 16.8% | 86.7% | 13.3% | 84.8% | 15.2% | 14.0% | 86.0% | 38.1% | 61.9% |
| Screening | 1.6% | 98.4% | 1.7% | 98.3% | 13.0% | 87.0% | 80.6% | 19.4% | 42.9% | 57.1% |
| Detect Change | 41.6% | 58.4% | 28.3% | 71.7% | 28.3% | 71.7% | 47.3% | 52.7% | 38.1% | 61.9% |
| Describe Status | 52.8% | 47.2% | 53.3% | 46.7% | 41.3% | 58.7% | 34.4% | 65.6% | 33.3% | 66.7% |

Note: Regular use of a measure was defined as being used with half or more than half of the last 40 children. Percentages show the percentage of SLPs who identified each purpose as a main purpose for which a measure was used. SLPs could select up to three main purposes for which each measure was used. On average, SLPs selected 2.7 purposes for each measure.

When the top three reasons for use of each language measure were combined, the most frequently identified reason for each of the three versions of the Clinical Evaluation of Language Fundamentals (CELF) core or language index subtests was presence of Australian norms. One third (61.6% or 77/125) of SLPs who regularly used the Clinical Evaluation of Language Fundamentals-4th Edition CELF-4 core or language index subtests selected this as a reason for doing so, 56.7% (37/60 SLPs) who regularly used the Clinical Evaluation of

Language Fundamentals-5th Edition (CELF-5) core or language index subtests selected this as a reason for doing so, and 56.5% (26/46 SLPs) who regularly used the Clinical Evaluation of Language Fundamentals: Preschool-2nd Edition (CELF:P-2) core or language index subtests selected this as a reason for doing so. Other frequently selected reasons for use of versions of the Clinical Evaluation of Language Fundamentals (CELF) measures included 'good psychometrics', 'employer requires use of assessment', and 'good for selecting goals'.).

The most frequently identified reason for use of the Renfrew Action Picture Test (RAPT) was quick to administer, with 91.5% (or 86/94 SLPs) who regularly used the Renfrew Action Picture Test (RAPT) selecting this as a reason doing so. Other frequently selected reasons for use of the Renfrew Action Picture Test (RAPT) included quick to score (72.3% or 68/94 SLPs), good for selecting goals (38.3% or 36/94 SLPs), and good for selecting intervention (29.8% or 28/94 SLPs). The most frequently identified reason for use of the Sutherland Phonological Awareness Test (SPAT) was good for selecting goals with 50.0% (or 21/42 SLPs) who regularly used the Sutherland Phonological Awareness Test (SPAT) selecting this as a reason doing so. Other frequently selected reasons for use of Sutherland Phonological Awareness Test (SPAT) included quick to score (42.9% or 18/42 SLPs), recently developed norms (42.9% or 18/42 SLPs), and quick to administer (38.1% or 16/42 SLPs). The reasons for which frequently used language measures were chosen for use are displayed in Table 6.4.

Table 6.4

*Reasons for Which Regularly Used Language Measures Were Chosen for Use*

| Reason | CELF-4 Core or Language Index (n=125) | | CELF-5 Core or Language Index (n=60) | | CELF:P Core Subtests (n=46) | | RAPT – Any Edition (n=94) | | SPAT – Any Edition (n=42) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Top reason | Not top reason | Top reason | Not top reason | Top reason | Not top reason | Top reason | Not top reason | Top reason | Not top reason |
| Australian Norms | **61.6%** | 38.4% | **56.7%** | 43.3% | **56.5%** | 43.5% | 4.3% | 95.7% | 23.8% | 76.2% |
| Good Psychometrics | **40.0%** | 60.0% | **36.7%** | 63.3% | **37.0%** | 63.0% | 2.1% | 97.9% | 11.9% | 88.1% |
| Employer requires use of assessment | **33.6%** | 66.4% | **38.3%** | 61.7% | 19.6% | 80.4% | 2.1% | 97.9% | 9.5% | 90.5% |
| Good for selecting goals | 32.0% | 68.0% | 36.7% | 63.3% | 34.8% | 65.2% | **38.3%** | 61.7% | **50.0%** | 50.0% |
| Referring Agent requires use of assessment | 24.0% | 76.0% | 13.3% | 86.7% | 8.7% | 91.3% | 1.1% | 98.9% | 0.0% | 100% |
| Only Available Assessment for purpose | 21.6% | 78.4% | 13.3% | 86.7% | 17.4% | 82.6% | 5.3% | 95.7% | 11.9% | 88.1% |
| Quick to administer | 16.8% | 83.2% | 5.0% | 95.0% | **39.1%** | 60.9% | **91.5%** | 8.5% | 38.1% | 61.9% |
| Good for selecting intervention | 15.2% | 84.8% | 13.3% | 86.7% | 21.7% | 78.3% | 29.8% | 70.2% | 9.5% | 90.5% |
| Only available Assessment for population | 12.8% | 87.2% | 1.7% | 98.3% | 13.0% | 87.0% | 2.1% | 97.9% | 9.5% | 90.5% |
| Quick to score | 12.8% | 87.2% | 11.7% | 88.3% | 21.7% | 78.3% | **72.3%** | 27.7% | **42.9%** | 57.1% |
| Good for selecting class strategies | 8.8% | 91.2% | 3.3% | 96.7% | 8.7% | 91.3% | 6.4% | 93.6% | 21.5% | 78.5% |
| Battery has reading/ writing | 6.0% | 94.0% | 8.3% | 91.7% | 4.3% | 95.7% | 0.0% | 100% | 2.4% | 97.6% |
| Battery has social abilities | 4.0% | 96.0% | 5.0% | 95.0% | 0.0% | 100% | 1.1% | 98.9% | 0.0% | 100% |
| Recently Developed Norms | 1.6% | 98.4% | 40.0% | 60.0% | 4.3% | 95.7% | 0.0% | 100% | **42.9%** | 57.1% |
| Inexpensive | 0.0% | 100% | 0.0% | 100% | 0.0% | 100% | 17.0% | 83% | 11.9% | 88.1% |
| Other Reason | 3.2% | 96.8% | 0.0% | 100% | 2.2% | 97.8% | 3.2% | 96.8% | 0.0% | 100% |

Note: Regular use was defined if the measure was used with half or more than half of the last 40 children.

Percentages show the percentage of SLPs who identified each reason as a top reason for use. SLPs could rank

up to three reasons for each measure. On average, each SLP selected 2.9 reasons for each measure. Bold font: indicates the three most frequently selected reasons for each measure.

## 6.5. Discussion

**6.5.1. Language measures and assessment procedures**. Findings from this study indicate that, although SLPs collectively use over 130 different language measures (i.e., measures with set guidelines for administration and scoring), on average each SLP used only seven different language measures with the last 40 children they assessed. This trend is similar to findings from previous surveys of SLP language assessment practice in the United States of America (Caesar & Kohler, 2009; Fulcher-Rood et al., 2018). The most regularly used language measures identified from the survey in this study were the core or language index subtests from versions of the Clinical Evaluation of Language Fundamentals (CELF), the Renfrew Action Picture Test (RAPT) and the Sutherland Phonological Awareness Test (SPAT). Aside from the Renfrew Action Picture Test (RAPT), all these measures have Australian normative data. The versions of the Clinical Evaluation of Language Fundamentals (CELF) also have normative data from the United States of America and United Kingdom and have also been reported in previous studies as being commonly used in these countries (Betz et al., 2013; Roulstone et al., 2015). The Renfrew Action Picture Test (RAPT) has normative data from the United Kingdom and has been identified as frequently used in the United Kingdom (Roulstone et al., 2015; Watson & Pennington, 2015). The most regularly used assessment procedures (i.e., procedures without set guidelines for administration and scoring) were interviews and observations. These procedures have also been identified in previous overseas studies as commonly used assessment procedures for school-aged children (Caesar & Kohler, 2009; Fulcher-Rood et al., 2018; Singh et al., 2016).

One different finding from this survey compared to previous surveys was low reported use of single-word vocabulary measures. Findings from previous surveys of SLP assessment practice in the United States of America have identified single word vocabulary

measures as being frequently used (Betz et al., 2013; Caesar & Kohler, 2009). This difference may be due to a lack of single word vocabulary measures with Australian norms, however, could also reflect a positive shift away from use of single word vocabulary measures by SLPs in general. Overall, findings from this current survey suggest the existence of broad trends in use of language measures and  assessment procedures  across English-speaking countries, with differences in use of language measures being potentially related to the availability of assessments with local normative data.

**6.5.2. Service methods.** Although research has not explicitly examined the practice of having other personnel conduct assessments, SLPs in this study reported using this as a service method. The inter-rater reliability between assessments conducted by SLPs and those conducted by other personnel is not known, therefore care should be taken in undertaking this practise until such time that these details are known. Therefore, SLPs should carefully consider the clinical implications of having other personnel conduct assessments, particularly with regards to the level of training that others may require to accurately conduct language assessment.

Findings from this study also indicated very limited use of ICTs as methods for conducting language assessment for primary school children. The use of ICTs has been shown to be a valid method for conducting language assessment (Waite et al., 2010) and has enormous potential to improve service accessibility for children and their families (O'Callaghan, McAllister, & Wilson, 2005). Use of ICTs may also reduce travel cost and save time for SLPs who provide outreach services (Mashima & Doarn, 2009). Therefore, further investigation is needed to examine reasons for low use of ICTs and identify actions to increase the utilisation of ICTs by SLPs as a method of conducting language assessment for primary school children.

The responses in this survey also indicate that, although interviews with parents and teachers were identified as the most regularly used assessment procedures, the majority of SLPs used these assessment procedures with less than half of the last 40 children they assessed. Furthermore, few proxy-reported language measures with set guidelines for administration and scoring were identified as regularly used by SLPs. These findings suggest that collecting information from parents and teachers may not be routine practice when SLPs assess the language abilities of primary school children. This finding is consistent with previous studies which have identified that although health professionals frequently spend time identifying family concerns, explaining assessment results or involving caregivers in interventions; they less frequently collect information from parents during the assessment process (Crais, Roy, & Free, 2006; Gillon et al., 2017). Barriers relating to successful SLP and teacher collaboration have also been identified in literature, which may impact on the frequency with which SLPs collect information from teachers during the assessment process (L. Wilson, McNeill, & Gillon, 2015). Given that both SLPs and families identify the importance of collecting assessment data from significant others (Crais et al., 2006), the is a need to increase the regularity with which SLPs use proxy-reported assessment methods to triangulate findings from language measures that are conducted by SLPs.

**6.5.3. Domains, purposes and reasons for use of language measures**. A notable finding in this study that the language measures reported by SLPs as being most regularly used all targeted semantics, morphosyntax and meta-abilities at word or sentence level. In addition, although the Clinical Evaluation of Language Fundamentals (CELF) core language and language index subtests were used regularly by many SLPs, few SLPs reported regularly using the Clinical Evaluation of Language Fundamentals (CELF) pragmatic profiles, Clinical Evaluation of Language Fundamentals (CELF-5) reading comprehension and structured writing tests or Clinical Evaluation of Language Fundamentals (CELF-5) pragmatic activities

checklist to assess social abilities and discourse. This finding suggests that SLPs prioritise word and sentence level tasks over discourse or text level tasks when assessing the language abilities of primary school children. As research identifies that discourse and text level tasks provide important information on a child's language abilities that may not be captured by word and sentence level measures (Lennox et al., 2018; Volden et al., 2017), there is a need to increase the regularity with which SLPs use language measures that go beyond word and sentence level.

Another notable finding in this study was that one quarter of SLPs reported that social abilities and discourse are areas they primarily focus on when using versions of the Clinical Evaluation of Language Fundamentals (CELF) core or language index subtests and the Renfrew Action Picture Test (RAPT), despite these measures not being designed for making judgments on these domains (Roulstone et al., 2015). These measures consist of structured table-top activities involving interaction with an adult and not allow for language abilities to be observed in a natural social context, which is important for making accurate judgements on social abilities and discourse (Trembath et al., 2016). Given these findings, it is important that SLPs take greater care to align the domains that measures are designed to assess, especially when interpreting assessment results to plan interventions.

Language measures that are used for diagnostic purposes should have evidence of psychometric quality (Dockrell & Marshall, 2015; Glover & Albers, 2007). The core or language index tests from versions of the  Clinical Evaluation of Language Fundamentals (CELF) have evidence of psychometric quality, therefore the finding in this survey that SLPs use these measures for diagnostic purposes is positive (Denman et al., 2017). Normative data from Australia also appeared to be an influencing factor for choice of diagnostic measures in this study. This is also a positive finding, although SLPs should be aware that this should not be the leading factor when selecting diagnostic measures as having normative data on the

population of interest does not necessarily mean a measure has sound reliability and validity (Denman et al., 2017).

It is also encouraging to note that SLPs identified Australian norms and good psychometrics as the most frequently reported reasons for choosing versions of the Clinical Evaluation of Language Fundamentals (CELF) core language and language index subtests, particularly given previous studies have identified that SLPs may not be considering psychometric quality when choosing language measures for use (Betz et al., 2013; Fulcher-Rood et al., 2018). The difference between findings from this study and previous studies could be a result of increased focus on psychometric quality in SLP literature in more recent times (Betz et al., 2013; Denman et al., 2017), although it may also be due to 'psychometric quality' being specifically listed as a response option in this study but not listed as a response option in previous studies (Betz et al., 2013).

Another frequently reported reason for use of versions of the Clinical Evaluation of Language Fundamentals (CELF) core language and language index subtests was 'employer requires use of assessment'. It seems likely that this reason may relate to requirements of providing norm-referenced scores from language measures for service eligibility (Fulcher-Rood et al., 2018). However, it could also reflect service agency policy that evidence-based language measures be used for diagnostic purposes. Further research is needed to develop a deeper understanding of the influence of service agency policy on language assessment practice (Fulcher-Rood et al., 2018).

Besides diagnostic purposes, SLPs in this study also reported using versions of the Clinical Evaluation of Language Fundamentals (CELF) core or language index subtests for purposes of predicting outcome, selecting intervention and detecting change. In addition, over 30% of SLPs identified 'good for selecting intervention goals' as one of the main reasons for choice of versions of the  Clinical Evaluation of Language Fundamentals (CELF) core or

language index subtests were chosen for use. Norm-referenced measures such as the Clinical Evaluation of Language Fundamentals (CELF) core or language index subtests are specifically designed to measure performance in relation to peers. These measures may not reflect performance in day-to-day life and may not assess each language target systematically or in enough depth to adequately determine which targets should be a focus for intervention (Ebert & Scott, 2014; Trembath et al., 2016). Similarly, these measures are not identified as being sensitive to changes in language ability over time (Bishop et al., 2016). As such, use of diagnostic measures such as the Clinical Evaluation of Language Fundamentals (CELF) core language and language index subtests for purposes of selecting intervention or detecting change does not align with evidence-based practice recommendations (Beck, 1995; Ebert & Scott, 2014; Huang et al., 1997).

In this survey, SLPs reported regularly using the Renfrew Action Picture Test (RAPT) for screening purposes. As the Renfrew Action Picture Test (RAPT) does not have evidence of diagnostic accuracy, it is not identified as being appropriate for screening language abilities (Glover & Albers, 2007; Renfrew, 2010). Furthermore, quality of psychometric properties was not a reason for choosing the Renfrew Action Picture Test (RAPT), suggesting that while SLPs may be aware of considering reliability and validity when choosing diagnostic measures, they may not give the same consideration when choosing screening measures.

The most frequently reported reasons for choice of the Renfrew Action Picture Test (RAPT) were 'quick to administer' and 'quick to score'. Limited time has been noted in a number of previous studies as a factor that may influence assessment practice (Fulcher-Rood et al., 2018; Pavelko et al., 2016; Roulstone et al., 2015; Westerveld & Claessen, 2014). This includes a previous study in the United Kingdom which also identified that the Renfrew Action Picture Test (RAPT) is chosen for use as it is quick to administer and score

(Roulstone et al., 2015). While cost of professional time needs to be considered, it is important that this is balanced with the selection of appropriate language measures. The findings from this survey indicate that SLPs need to give greater consideration to the features that make different language measures suitable for different purposes when choosing measures for use (Wade, 2004).

      **6.5.4. Implications.** While the practice of having other personnel conducting assessments has the potential to alleviate time pressures experienced by SLPs, risks of misdiagnosis may be present if training in conducting language assessment is not included in the job training of other personnel. Furthermore, test publishers may stipulate that those purchasing and conducting a particular assessment have specific qualifications (Pearson Education, 2018). Until further research is conducted to establish outcome of having other personnel conducting assessments, SLPs should be aware that they are legally and ethically responsible for services conducted by other personnel on their behalf and, as a result, exercise caution when engaging in this practice (American Psychological Association, 2000; American Speech and Hearing Association, 1997-2018; Speech Pathology Australia, 2016). There is some literature identifying that other personnel may be appropriately trained to support delivery of language interventions for certain groups of children (Boyle et al., 2007). Therefore, SLPs may be well-advised to consider utilising the time of other personnel, such as teacher-aides and therapy assistants, to support delivery of some specific interventions rather than conducting assessments.

      Limited use of ICTs for delivering assessment services may mean that many children and their families will continue to experience limitations with access to assessment services due to long travel distances, high costs and limited choice in service providers (O'Callaghan et al., 2005). Underutilisation of language measures that target social-abilities and discourse and underutilisation of proxy-reported methods may mean that functional abilities are not

adequately assessed and that intervention goals developed from assessment data are not representative of, or indeed targeting, a child's performance in everyday communication contexts (Ebert & Scott, 2014; Kover, Davidson, Sindberg, & Weismer, 2014; Tager-Flusberg et al., 2009). In addition, using measures for purposes that the measures were not designed for may compromise the accuracy of decisions made from assessment results. For example, the use of screening measures that do not have evidence of reliability, validity and diagnostic accuracy may mean that children are frequently misclassified as not having a disorder when they do and vice versa (Dockrell & Marshall, 2015).

   **6.5.5. Future directions**. Findings from this study identify that actions need to be taken to improve the decisions SLPs make when assessing the language abilities of primary school-aged children. The regularity with which SLPs use ICTs and proxy-reported methods needs to increase; therefore, further research is needed to examine the barriers and facilitators that influence regular use of these assessment delivery methods. To build SLP capacity with regards to choosing language measures that are appropriate for the domains being assessed and the purposes for which assessment data are to be used, it is recommended that both undergraduate training and post-graduate continuing professional development for SLPs places greater emphasis on the particular domains and purposes for which different language measures are suitable for (Wade, 2004).

   Findings from this survey also identified that a vast array of  language measures are in use for assessing the language abilities of school-aged children. Having such a large array of options may be over-whelming, thus making it difficult for SLPs to make sound decisions about which measures to use (McCabe, 2018). The future development of clinical practice guidelines or decision-making aids may assist SLPs to make evidence-based decisions when assessing the language abilities of primary school children.

**6.5.6. Limitations.** This study has a number of strengths, notably with regards to provision of agreed-upon definitions of terms for describing assessment methods, domains and purposes; and the presence of a scale with numeric qualifiers for rating frequency with which measures and methods are used (Blais & Grondin, 2011). Nonetheless, as with all surveys, this study relies on reported information. It is possible that some participants may not have accurately identified all the language measures they used. There were also a proportion of language measures that could not be identified from participant descriptions and were thus not able to be included.

Due the need to keep the survey length manageable for participants, SLPs in this survey only provided information on the domains assessed, purposes of use, and reasons for choice of use in relation to language measures. Therefore, this study does not provide information on the choices SLPs make with regards to assessment procedures. This survey also did not examine the combination of language measures and assessment procedures that are used for each child and, therefore, does not provide information relevant to management of individual cases. In addition, data was not collected on the types of assessments that are conducted by other personnel or the purposes and reasons for which other personnel conduct assessments. This information would assist in building a more in-depth understanding of the extent to which this assessment method is used.

**6.6. Conclusion**

This study investigated the language measures (i.e., measures with set guidelines for administration and scoring), assessment procedures (i.e., procedures without set guidelines for administration and scoring) and assessment delivery methods that Australian SLPs use when assessing the language abilities of primary school children. Collectively, SLPs used a large array of language measures, although each SLP used only a small number of measures regularly. SLPs reported that other personnel are involved in conducting language

assessment, despite limited literature to support this practice. Some assessment methods may be underutilised by SLPs, such as assessments conducted by ICTs and proxy-reported assessment methods. SLPs appear to select diagnostic measures based on psychometric properties, but not screening measures. In addition, SLPs may be using measures to target domains that measures are not ideally suited to measuring. Overall, these findings identify the need for greater emphasis to be placed on evidence-based practice when SLPs choose language measures, assessment procedures and assessment delivery methods for use with primary school children.

**References for Chapter 6**

American Psychological Association. (2000). *Report of the task force on test user qualifications*. Retrieved from

https://www.apa.org/science/programs/testing/qualifications.pdf

American Speech and Hearing Association. (1997-2018). *Speech language pathology assistants*. Retrieved from https://www.asha.org/Practice-Portal/Professional-Issues/Speech-Language-Pathology-Assistants/

Beck, A. R. (1995). Language assessment methods for three age groups of children. *Journal of Communication Development, 17*(2), 51-56. doi:10.1177/152574019501700206

Betz, S. K., Eickhoff, J. K., & Sullivan, S. F. (2013). Factors influencing the selection of standardized tests for the diagnosis of Specific Language Impairment. *Language, Speech and Hearing Services in Schools, 44*, 133-143. doi:10.1044/0161-1461(2012/12-0093)

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS ONE, 11*(7), e0158753. doi:10.1371/journal.pone.0158753

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE-2 consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*. doi:10.1111/jcpp.12721

Blais, J. G., & Grondin, J. (2011). The influence of labels associated with anchor points of likert-type response scales in survey questionnaires. *Journal of Applied Measurement, 12*, 370–386.

Boyle, J., McCartney, E., Forbes, J., & O'Hare, A. (2007). A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment. *Health Technology Assessment, 11*(25), iii-iv, xi.

Caesar, L. G., & Kohler, P. D. (2009). Tools clinicians use: A survey of language assessment procedures used by school-based speech-pathologists. *Communication Disorders Quarterly, 30*(4), 226-236. doi:10.1177/1525740108326334

Crais, E. R., Roy, V. P., & Free, K. (2006). Parents' and professionals' perceptions of the implementation of family-centered practices in child assessments. *American Journal of Speech-Language Pathology, 15*(4), 365-377. doi:10.1044/1058-0360(2006/034)

Denman, D., Cordier, R., Kim, J.-H., Munro, N., & Speyer, R. (Under Review). Language Assessment for Primary School Children (Part I): What Factors Influence Assessment Use? .

Denman, D., Kim, J.-H., Munro, N., Speyer, R., & Cordier, R. (2019). Describing language assessments for school-aged children: A Delphi study. *International Journal of Speech Language Pathology*, Advance Online Publication. doi:10.1080/17549507.2018.1552716

Denman, D., Speyer, R., Munro, N., Pearce, W., Chen, Y., & Cordier, R. (2017). Psychometric properties of language assessments for children aged 4-12 years: A systematic review. *Frontiers in Psychology, 8*, 1-28. doi:10.3389/fpsyg.2017.01515

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. *Child and Adolescent Mental Health, 20*(2), 116-125. doi:10.1111/camh.12072

Eadie, P. (2003). Speech pathology assessment practices: One assessment or many? *Advances in Speech Language Pathology, 5*(1), 65-68. doi:10.1080/14417040510001669081

Ebert, K. D., & Scott, C. M. (2014). Relationships between narrative language samples and norm-referenced test scores in language assessments of school-age children. *Language, Speech, and Hearing Services in Schools, 45*, 337–350. doi:10.1044/2014_LSHSS-14-0034

Fulcher-Rood, K., Castilla-Earls, A. P., & Higginbotham, J. (2018). School-based speech-language pathologists' perspectives on diagnostic decision making. *American Journal of Speech-Language Pathology, 27*(2), 796-812. doi:10.1044/2018_AJSLP-16-0121

Gillon, G., Hyter, Y., Fernandes, F. D., Ferman, S., Hus, Y., Petinou, K., . . . Westerveld, M. (2017). International survey of speech-language pathologists' practices in working with children with autism spectrum disorder. *Folia Phoniatrica et Logopaedica, 69*(1-2), 8-19. doi:10.1159/000479063

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117-135. doi:10.1016/j.jsp.2006.05.005

Harrison, L. J., McLeod, S., Berthelsen, D., & Walker, S. (2009). Literacy, numeracy, and learning in school-aged children identified as having speech and language impairment in early childhood. *International Journal of Speech-Language Pathology, 11*(5), 392-403. doi:10.1080/17549500903093749

Hegde, M. N., & Pomaville, F. (2017). Assessment of communication disorders in children. In *Assessment of communication disorders in children: Resources and protocols* (pp. 3-16). Sa Diego, CA: Plural Publishing.

Huang, R. J., Hopkins, J., & Nippold, M. A. (1997). Satisfaction with standardized language testing: A survey of speech-language pathologists. *Language, Speech, and Hearing Services in Schools, 28*(1), 12-29. doi:10.1044/0161-1461.2801.12

IBM Corp. (Released 2011). IBM SPSS Statistics for Windows Version 20.0. Armonk, NY: IBM Corp.

Kover, S. T., Davidson, M. M., Sindberg, H. A., & Weismer, S. E. (2014). Use of the ADOS for assessing spontaneous expressive language in young children with ASD: A comparison of sampling contexts. *Journal of Speech, Language, and Hearing Research, 57*(6), 2221-2233. doi:10.1044/2014_JSLHR-L-13-0330

Lennox, M., Westerveld, M. F., & Trembath, D. (2018). Should we use sentence-or text-level tasks to measure oral language proficiency in year-one students following whole-class intervention? *Folia Phoniatrica et Logopaedica, 69*(4), 169-179. doi:10.1159/000485974

Letts, C., Edward, S., Schaefer, B., & Sinka, I. (2014). The New Reynell Developmental Language Scales: Descriptive account and illustrative case study. *Child Language Teaching and Therapy, 30*(1), 103-116. doi:10.1177/0265659013492784

Mashima, P. A., & Doarn, C. R. (2009). Overview of telehealth activities in speech-language pathology. *Telemedicine and e-Health, 14*(10), 1101-1117. doi:10.1089/tmj.2008.0080

McCabe, P. J. (2018). Elizabeth Usher Memorial Lecture: How do we change our profession? Using the lens of behavioural economics to improve evidence-based practice in speech-language pathology. *International Journal of Speech-Language Pathology, 20*(3), 300-309. doi:10.1080/17549507.2018.1460526

Neilsen, R. (2003). *Sutherland Phonological Awareness Test - Revised*. Camberwell, Australia: ACER.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., . . . Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of

language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry, 57*(11), 1247-1257. doi:10.1111/jcpp.12573

O'Callaghan, A. M., McAllister, L., & Wilson, L. (2005). Barriers to accessing rural paediatric speech pathology services: Health care consumers' perspectives. *Australian Journal of Rural Health, 13*(3), 162-171. doi:10.1111/j.1440-1854.2005.00686.x

Pavelko, S. L., Owens, R. E., Ireland, M., & Hahs-Vaughn, D. L. (2016). Use of language sample analysis by school-based SLPs: Results of a nationwide survey. *Language, Speech, and Hearing Services in Schools, 47*(3), 246-258. doi:10.1044/2016_LSHSS-15-0044

Pearson Education. (2018). PreK-16 education and special needs: Qualifications policy. Retrieved from https://www.pearsonclinical.com/education/qualifications.html

Polit, D. F. (2015). Assessing measurement in health: Beyond reliability and validity. *International Journal of Nursing Studies, 52*, 1746-1753. doi:10.1016/j.ijnurstu.2015.07.002

Renfrew, C. (2010). *Renfrew Action Picture Test*. Camberwell, Australia: ACER.

Roulstone, S. (2015). Exploring the relationship between client perspectives, clinical expertise and research evidence. *International Journal of SpeechLanguage Pathology, 17*(3), 211-222. doi:10.3109/17549507.2015.1016112

Roulstone, S., Marshall, J., Powell, G. G., Goldbart, J., Wren, Y. E., Coad, J., . . . Coad, R. A. (2015). Evidence-based intervention for preschool children with primary speech and language impairments: Child Talk – an exploratory mixed-methods study. *Programme Grants for Applied Research, 3*(5), 1-408. doi:10.3310/pgfar03050

Singh, J. S., Chan, M. Y., & Rusli, Y. A. (2016). Practise patterns of Malaysian speech-language pathologists in managing children with speech and language delay/disorder.

*International Journal of Speech-Language Pathology, 18*(6), 560-570.
doi:10.3109/17549507.2016.1139624

Speech Pathology Australia. (2016). *Position paper: Allied health assistants - speech pathology*. Retrieved from
http://www.speechpathologyaustralia.org.au/SPAweb/Resources_for_Speech_Patholo gists/Professional_Resources/HTML/Allied_Health_Assistants.aspx?WebsiteKey=fc2 020cb-520d-405b-af30-fc7f70f848db

Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., . . . Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research, 52*(3), 643-652. doi:10.1044/1092-4388(2009/08-0136)

Trembath, D., Westerveld, M., & Shellshear, L. (2016). Assessing spoken language outcomes in children with ASD: A systematic review. *Current developmental Disorders Reports, 3*(33), 33-45. doi:10.1007/s40474-016-0068-8

Tucker, J. K. (2012). Perspectives of speech-language pathologists on the use of telepractice in schools: Quantitative survey results. *International Journal of Telerehabilitation, 4*(2), 61-72. doi:10.5195/ijt.2012.6100

Volden, J., Dodd, E., Engel, K., Smith, I. M., Szatmari, P., Fombonne, E., & Vaillancourt, T. (2017). Beyond sentences: Using the Expression, Reception, and Recall of Narratives Instrument to assess communication in school-aged children with Autism Spectrum Disorder. *Journal of Speech, Language, and Hearing Research, 60*(8), 2228-2240. doi:10.1044/2017_JSLHR-L-16-0168

Wade, D. T. (2004). Assessment, measurement and data collection tools. *Clinical Rehabilitation, 18*, 233-237. doi:10.1191/0269215504cr183ed

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010). Internet-based telehealth assessment of language using the CELF–4. *Language Speech and Hearing Services in Schools, 41*, 445-458. doi:10.1044/0161-1461(2009/08-0131)

Westerveld, M. F., & Claessen, M. (2014). Clinician survey of language sampling practices in Australia. *International Journal of Speech-Language Pathology, 16*(3), 242-249. doi:10.3109/17549507.2013.871336

Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals: Preschool - 2nd Edition* (2 ed.). Bloomington, MN: Pearson.

Wiig, E. H., Semel, E., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals - 4th Edition* (5 ed.). Bloomington, MN: Pearson.

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals - 5th Edition*. Bloomington, MN: Pearson.

Williams, C., & McLeod, S. (2012). Speech-language pathologists assessment and intervention practices with multi-lingual children. *International Journal of Speech-Language Pathology, 14*(3), 292-305. doi:10.3109/17549507.2011.636071

Wilson, K. S., Blackmon, R. C., Hall, R. E., & Elcholtz, G. E. (1991). Methods of language assessment: A survey of California public school clinicians. *Language, Speech and Hearing Services in Schools, 22*, 236-241. doi:10.1044/0161-1461.2204.236

Wilson, L., McNeill, B., & Gillon, G. T. (2015). The knowledge and perceptions of prospective teachers and speech language therapists in collaborative language and literacy instruction. *Autism & Developmental Language Impairments, 2*, 1-17. doi:10.1177/0265659015585374

**Supplementary Appendix 6.1.**

Survey for Australian Speech-Language Pathologists (Questions Relating to Part II)


**I consent to answering questions in an on-line survey and for my responses to be used for the purposes described above**.

*[Yes/No response. If no, skip to end of survey]*

**SECTION 1**

**Do you have current practicing membership with Speech Pathology Australia?**

*[Yes/No response]*

**Please indicate your age:**

*[Multiple choice response]*

**Please indicate your gender:**

*Multiple choice responses*

**Is English your first language?**

*[Yes/No response]*

**Please indicate the number of years since you graduated as a speech pathologist:**

*[Multiple choice responses]*

**Are you currently in paid employment as an SLP?**

(This may include non-clinical roles in the field of SLP)

*[Yes/No response. If no, skip to end of survey]*

**Please give your postcode**

If you are currently employed please give your work postcode.

If you work in more than one job, please give the postcode the job in which you work the most hours or if hours are equal, choose the job in which you have worked the longest.

If you are not currently employed, please give your home postcode (e.g. full-time student or

are on extended leave)

*[Open text response]*

**Please indicate the amount of time per week you are currently employed** (paid to work)

**as a speech pathologist** (across all jobs)

*[Multiple choice response]*

**Please select the box/es that best describe/s your current paid work as a speech**

**pathologist** (**across all jobs**).

Select all appropriate options e.g. if you provide services to children aged 0-6 years then

select both of the first two options, or if you provide services to clients of all ages then select

all the clinical service provision boxes.

*[Multiple choice response. If participant does not select response "work clinically with*

*children 4-12 years]*

**In the last 12 months, have you provided clinical services to at least 40 children aged 4-**

**12 years with oral or written language disorders?** (Includes supervision of SLP students

who provide clinical services)

For the purposes of this survey, the term 'children with language disorder' is used broadly to

refer to any children who require support for oral or written language (i.e. semantics, syntax,

morphology, phonemic awareness, discourse or social abilities), regardless of the primary

diagnosis, aetiology or co morbidities associated with the language support needs.

Children who have a lack of familiarity with standard Australian English are also included in

this group if they are accessing your services for language assessment.

This includes children with: developmental language disorder, dyslexia, autism spectrum

disorder, learning difficulties, intellectual disability or language disorder associated with

conditions such as traumatic brain injury or hearing impairment.

*[Yes/No response. If no, skip to end of survey]*

**SECTION 2**

**Indicate the option that best describes the agency through which you provide clinical services to children aged 4-12 years with language disorder:**

If you work in more than one job, please complete this survey for the job through which you most frequently service children aged 4-12 years with language disorder.

*[Multiple choice responses]*

**Consider the last 40 children (4-12 years with language disorder) who accessed your services. How many children <u>did not</u> have standard Australian English as their first or only language?** (i.e. how many children were bi-lingual, learning English as a second language or were considered as having a cultural/linguistic difference)

*[Likert scale response]*

**SECTION 3**

**In this section of the survey you are asked to estimate the frequency in which different types of oral and written assessments were used** (considering the last 40 children who accessed your services for assessment).

Terminology may be used differently across the profession, therefore it is important that you read the survey information carefully and select answers based on how terms are defined in this survey (and not how you define them or have seen them defined elsewhere).

When answering the questions, please consider:

All assessment data gathering activities (e.g. tests, observations, parent or teacher interviews and collection of case histories)

Assessments conducted for any purpose (e.g. diagnosis, screening, detecting change, selecting intervention**)**

**Assessments conducted by other trained professionals on your behalf**

(i.e. not conducted by you as an SLP or by an SLP student supervised by you)

Language assessments conducted by another person (through testing or language sampling) as part of your SLP service provision. You (or SLP student) may assist by providing training or guidance, however the other person has the primary role in conducting the assessment

Examples include:

Language screening tests conducted by teacher-aides or therapy assistants with results used to inform your SLP service provision

Language assessments conducted by specialist teachers (e.g. as part of a school support team assessment process) with results used to inform your SLP service provision

Language assessments conducted by an OT or a psychologist (e.g. during a multi-disciplinary team assessment) with results used to inform your SLP service provision

*[Likert scale response]*

**Please list the other professional/s and the assessments they conducted**

For example:

Teacher-aides conducted language screening prior to SLP assessment

Special Education Teacher administered PLS-5 as part of multidisciplinary team assessment

*[Open text response*]

**Assessment using Information & Communication Technologies (ICTs)**

Assessments conducted with the assessor and the student communicating through ICTs (telehealth).

Examples include: Skype, video-conferencing, web-conferencing, telephone.

Note: Technology that is not used for two-way communication (e.g. audio/video recording devices, Boardmaker software or scoring assistance software) is not considered use of ICTs.

*Likert scale response]*

**SECTION 4**

**In this section of the survey you are asked specifically about the names of the oral and written language assessments that were used by you or by others on your behalf**

(considering the last 40 children assessed)

**What are the main purposes for which the [standardised assessment listed by participant] was used?**

You may select up to three options.

Please read the definition of each term carefully to ensure that you make the correct selection.

*[Multiple choice response]*

- *Predict outcome/planning: Predict need for intervention, identify classroom support needs, identify type of curriculum differentiation or predict risk of poor future outcome*

- *Select intervention: Identify a suitable intervention approach or select intervention goals/targets*

- *Plan dosage: Predict amount or intensity (dosage) of intervention*

- *Screening: Identify children who may have a disorder and require further diagnostic assessment (i.e. identify if diagnostic assessment should be conducted and/or the domains to be targeted in diagnostic assessment)*

- *Diagnostic: Diagnose a condition or make a comparison with peers (i.e. identify the presence or severity of a disorder, or determine if functioning is different to peers)*

- *Detect change: Measure change in status, measure outcomes or monitor progress over time*

- *Describe status: Assessment for the purpose of profiling strengths and weaknesses or describing and explaining a particular aspect of a child's functioning*

**What domains do you primarily focus on when you use the [standardised assessment listed by participant]?**

Please read definitions carefully before selecting your answers.

*[Multiple choice response]*

- *Semantics: Understanding and expression of words and word meanings (e.g. vocabulary, word retrieval, lexical meaning)*

- *Morphosyntax: Understanding and expression of different word forms and the order and combination of words in sentences (i.e. morphology, word order, sentence construction)*

- *Social Abilities & Discourse: Giving and making meaning in social context or communication for social purposes. Includes:*

- *Pre-linguistic communication (e.g. gestures, joint attention)*

   *Range of communicative intentions/purposes*

   *Non-verbal communication (e.g. understanding facial expressions)*

   *Non-literal language (e.g. idioms, sarcasm)*

   *Matching communication style to social context*

   *Conversation conventions (e.g. topic selection and maintenance)*

   *Text cohesion (e.g. verbal fluency)*

   *Text organization (e.g. story grammar, episodic structure, sequencing of information)*

- *Meta-Abilities: Ability to think about own thought processes and understand how to regulate these processes for effective learning. Includes:*

   *Meta-cognition: knowledge and use of strategies for managing and self-monitoring own learning*

- *Meta-language: phonological awareness, spelling conventions, meta-linguistic and meta-narrative skills*

   *Meta-pragmatics: awareness and application of social skills or conventions*

- *Executive Functioning: Cognitive processes necessary for execution of goal-directed, controlled, purposeful behavior. Includes:*

  *Inhibition (self-control): Suppression of inappropriate thoughts, comments and behaviors in order to focus and attend to tasks*

  *Emotion control (self-regulation): Ability to manage emotions for goal achievement and task completion.*

  *Working memory: Retention, processing and manipulation of pieces of information for short periods of time in order to complete required tasks*

  *Organisation (strategic planning): Organisational strategies for task completion*

  *Mental flexibility: Integration of prior knowledge and experiences when completing tasks and effective application of different rules for different situations*

  *Sustained attention: Ability to maintain attention to tasks despite distractions and fatigue*

**What are the main reasons that you use [standardised assessment listed by participant] over other assessments?** Rank (by placing numbers in the boxes) up to 3 reasons, with 1 being the biggest or most influential reason.

*[Rank option response with open text boxes for options not listed]*

- *Employer requires use of this assessment (e.g. to determine eligibility for funding)*

- *Referring agent requires use of this assessment (e.g. to determine eligibility for funding)*

- *Is the only assessment available in my workplace for the purpose/s I use it for*

- *Is the only assessment available in my workplace for the client population I use it for*

- *Was (relatively) inexpensive to purchase compared to other similar assessments*

- *Is (relatively) quick and easy to administer*

- *Is (relatively) quick and easy to score/analyze results*

- *The assessment battery has tests for reading and writing as well as oral language*

- *The assessment battery includes assessments for social abilities/pragmatics*

- *Is recently published or has recently developed norms*

- *Has Australian normative data*

- *Has sound reliability, validity and diagnostic accuracy*

- *Provides useful information for selecting intervention approach*

- *Provides useful information for identifying or measuring intervention goals or targets*

- *Provides useful information for identifying classroom strategies or supports*

- *Other reason. Please specify:*

**Supplementary Appendix 6.2.**

Language Measures used by SLPs and Regularity of Use (n=335).

| Assessment | Author and publication year | SLPs who used assessment | SLPs who used assessment regularly* |
|---|---|---|---|
| CELF-4 Core and Language Index Subtests (Clinical Evaluation of Language Fundamentals - Fourth Edition) | Wiig, Semel, and Secord (2004) | 245 (73.1%) | 125 (37.3%) |
| CELF:P Core Language Subtests (Clinical Evaluation of Language Fundamentals Preschool – Any Edition) | Wiig, Secord, and Semel (2004) | 222 (66.2%) | 46 (13.7%) |
| RAPT (Renfrew Action Picture Test – Any Edition) | Renfrew (2010a) | 214 (63.8%) | 93 (27.8%) |
| CELF-5 Core and Language Index Subtests (Clinical Evaluation of Language Fundamentals - Fifth Edition) | Wiig, Semel, and Secord (2013) | 191 (57.0%) | 60 (17.9%) |
| [a] SPAT (Sutherland Phonological Awareness Assessment – Any Edition) | Neilsen (2003) | 191 (57.0%) | 42 (12.5%) |
| PLS (Preschool Language Scales – Any Edition) | Zimmerman, Steiner, and Pond (2011) | 138 (41.1%) | 8 (2.4%) |
| CELF-4 Working Memory Subtests (Clinical Evaluation of Language Fundamentals - Fourth Edition) | Wiig, Semel, et al. (2004) | 65 (19.4%) | 7 (2.1%) |
| [b] TNL (Test of Narrative Language) | Gillam and Pearson (2004) | 57 (17.0%) | 5 (14.9%) |
| [ab] CELF-5 Reading Comprehension and Structured Writing Subtests (Clinical Evaluation of Language Fundamentals - Fifth Edition) | Wiig et al. (2013) | 52 (15.2%) | 3 (0.1%) |
| [bc] CELF-5 Pragmatic Profile (Clinical Evaluation of Language Fundamentals - Fifth Edition) | Wiig et al. (2013) | 51 (15.2%) | 7 (2.1%) |
| TOPS (Test of Problem Solving) (Elementary or Adolescent – Any Edition) | Bowers, Huisingh, and LoGiudice (2005) | 47 (14.1%) | 3 (0.1%) |
| [bc] CELF-4 Pragmatics Profile (Clinical Evaluation of Language Fundamentals - Fourth Edition) | Wiig, Semel, et al. (2004) | 36 (10.7%) | 3 (0.1%) |
| [ab] YARC - Passage Reading (York Assessment of Reading for Comprehension – Any Edition) | Snowling et al. (2009) | 35 (10.4%) | 7 (2.1%) |

| Assessment | Author and publication year | SLPs who used assessment | SLPs who used assessment regularly* |
|---|---|---|---|
| [b] RBS (Renfrew Bus Story – Any Edition) | Renfrew (2010b) | 33 (9.9%) | 8 (2.4%) |
| [c] CCC (Children's Communication Checklist – Any Edition) | Bishop (2003) | 32 (9.5%) | 4 (1.2%) |
| [ab] NARA (Neale Analysis of Reading Ability – Any Edition) | Neale (1999) | 32 (9.5%) | 3 (0.1%) |
| PPVT (Peabody Picture Vocabulary Test – Any Edition) | Dunn and Dunn (2007) | 31 (9.3%) | 4 (1.2%) |
| Reynell (Reynell Developmental Scales – Any Edition) | Edwards, Letts, and Sinka (2011) | 28 (8.45) | 1 (0.3%) |
| [a] CTOPP (Comprehensive Test of Phonological Processing – Any Edition) | Wagner, Torgesen, Rashotte, and Pearson (1999) | 24 (7.2%) | 7 (2.1%) |
| [a] QUIL (Queensland Inventory of Literacy) | Dodd and The University of Queensland Dept. of Speech Pathology & Audiology (1996) | 22 (6.6%) | 0 (0.0%) |
| [a] YARC - Early Reading (York Assessment of Reading for Comprehension) | Snowling et al. (2009) | 22 (6.6%) | 1 (0.3%) |
| CELF-4 Screening Test (Clinical Evaluation of Language Fundamentals Screening Test- Fourth Edition) | Semel, Wiig, and Secord (2006) | 21 (6.3%) | 7 (2.1%) |
| [b] TNL -2 (Test of Narrative Language – Second Edition) | Gillam and Pearson (2017) | 21 (6.3%) | 1 (0.3%) |
| [c] Communication Matrix | Rowland and Fried-Oken (2004) | 20 (6.0%) | 1 (0.3%) |
| TOLD - I (Test of Language Development – Intermediate – Any Edition) | Newcomer and Hammill (2008) | 19 (5.7%) | 0 (0.0%) |
| CASL (Comprehensive Assessment of Spoken Language) | Carrow-Woolfolk (2017) | 18 (5.4%) | 2 (0.6%) |
| [b] ERRNI (Expression, Reception, Recall of Narrative Instrument) | Bishop (2004) | 16 (4.8%) | 0 (0.0%) |
| OWLS-II (Oral and Written Language Scales - 2nd Edition) Listening and/or Speaking components | Carrow-Woolfolk (2011) | 15 (4.5%) | 0 (0.0%) |
| TOLD - P (Test of Language Development – Primary – Any Edition) | Hammill and Newcomer (2008) | 15 (4.5%) | 0 (0.0%) |

| Assessment | Author and publication year | SLPs who used assessment | SLPs who used assessment regularly* |
|---|---|---|---|
| Bureau Test of Auditory Comprehension | Health Commission of New South Wales (1990) | 15 (4.5%) | 3 (0.1%) |
| [b] Peter and the Cat Retell | Leitao and Allan (2003) | 13 (3.8%) | 1 (0.3%) |
| PLAI (Preschool Language Assessment Instrument) | Blank, Rose, and Berlin (2003) | 13 (3.8%) | 1 (0.3%) |
| TOPL-2 (Test of Pragmatic Language - Second edition) | Phelps-Terasaki and Phelps-Gunn (2007) | 13 (3.8%) | 0 (0.0%) |
| [bc] CELF:P (Clinical Evaluation of Language Fundamentals: Preschool- Second Edition) Pragmatics Profile | (Wiig, Secord, et al., 2004) | 11 (3.3%) | 1 (0.3%) |
| [ab] OWLS-II (Oral and Written Language Scales - 2nd Edition) Reading and Writing Tests | Carrow-Woolfolk (2011) | 10 (3.0%) | 1 (0.3%) |
| [a] YARC (York Assessment of Reading for Comprehension) – Not otherwise specified | Snowling et al. (2009) | 10 (3.0%) | 0 (0.0%) |
| [b] ONAP (Oral Language Assessment Package): DECD South Australia | Government of South Australia: Department of Education and Children's Services | 9 (2.7%) | 1 (0.3%) |
| [a] South Australian Spelling Test | Westwood (2005) | 9 (2.7%) | 1 (0.3%) |

Note: Language measures that were listed by 2.4% (8/335) or less participants are not included in this table

*Regular use was identified if a measure was used with 20 or more of the last 40 children who received services (i.e. half or more than half of children)

[a] Language measure that targets written language

[b] Language measure that targets social abilities and discourse

[c] proxy-reported assessment method

**References for Supplementary Appendix 6.2.**

Bishop, D. V. M. (2003). *Children's Communication Checklist - Second Edition*. London, England: Psycholgical Corporation.

Bishop, D. V. M. (2004). *Expression, Reception and Recall of Narrative Instrument*. US: Pearson.

Blank, M., Rose, S. A., & Berlin, L. J. (2003). *Preschool Language Assessment Instrument (PLAI-2)*. Austin, TX: ProEd.

Bowers, L., Huisingh, R., & LoGiudice, C. (2005). *Test of Problem Solving 3rd Edition (TOPS-3)*. East Moline, IL: LinguiSystems, Inc.

Carrow-Woolfolk, E. (2011). *Oral and Written Language Scales - Second Edition* (2 ed.). Bloomington, MN: Pearson.

Carrow-Woolfolk, E. (2017). *Comprehensive Assessment of Spoken Language - Second Edition*. Torrance, CA: WPS.

Dodd, B., & The University of Queensland Dept. of Speech Pathology & Audiology. (1996). *Queensland University Inventory of Literacy (QUIL)*. Brisbane, Australia: The University of Queensland Dept. of Speech Pathology & Audiology.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test - Fourth Edition* (Fourth Edition ed.). Bloomington, MN: Pearson.

Edwards, S., Letts, C., & Sinka, I. (2011). *New Reynell Developmental Language Scales*. London, England: GL Assessment.

Gillam, R. B., & Pearson, N. A. (2004). *Test of Narrative Language (TNL)*. Austin: TX: PRO-Ed Inc.

Gillam, R. B., & Pearson, N. A. (2017). *Test of Narrative Language: Second Edition*. Austin: TX: PRO-Ed Inc.

Government of South Australia: Department of Education and Children's Services. *Oral Narrative Assessment Package (ONAP)*. Adelaide, Australia: Government of South Australia: Department of Education and Children's Services.

Hammill, D. D., & Newcomer, P. L. (2008). *Test of Oral Language Development - Primary: 4th Edition*. Austin, TX: Pro-Ed.

Health Commission of New South Wales. (1990). *The Bureau Auditory Comprehension Test*. Sydney, Australia: Health Commission of New South Wales.

Leitao, S., & Allan, L. (2003). *Peter and the Cat Narrative Assessment*: Black Sheep Press Ltd.

Neale, M. D. (1999). *Neale Analysis of Reading - Third Edition*. Camberwell, Australia: ACER.

Neilsen, R. (2003). *Sutherland Phonological Awareness Test - Revised*. Camberwell, Australia: ACER.

Newcomer, P. L., & Hammill, D. D. (2008). *Test of Language Development - Intermediate: 4th Edition* (4 ed.). Austin, TX, U.S.: Pro-Ed.

Phelps-Terasaki, D., & Phelps-Gunn, T. (2007). *Test of Pragmatic Language - 2nd Edition (TOPL-2)*. Torrance, CA: WPS.

Renfrew, C. (2010a). *Renfrew Action Picture Test*. Camberwell, Australia: ACER.

Renfrew, C. (2010b). *Renfrew Bus Story*. Camberwell, Australia: ACER.

Rowland, C., & Fried-Oken, M. (2004). Communication Matrix. Retrieved from https://communicationmatrix.org/

Semel, E., Wiig, E., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals - Fourth Edition Screening Test (CELF-4 Screening)*. Bloomington, MN: Pearson.

Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., & Nation, K. (2009). *York Assessment of Reading Comprehension (YARC)*. London, England: GL Assessment.

Wagner, R., Torgesen, J., Rashotte, C., & Pearson, N. A. (1999). *Comprehensive Test of Phonological Processing–Second Edition*. Austin, TX: ProEd.

Westwood, P. (2005). *Spelling: Approaches to teaching and assessment*: Australian Council for Educational Research.

Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals: Preschool - 2nd Edition* (2 ed.). Bloomington, MN: Pearson.

Wiig, E. H., Semel, E., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals - 4th Edition* (5 ed.). Bloomington, MN: Pearson.

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals - 5th Edition*. Bloomington, MN: Pearson.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales-5th Edition* (5 ed.). Bloomington, MN: Pearson.

**Chapter 7.**

**Summary, Discussion and Conclusions**

The research reported in this thesis contributed important knowledge in three areas of research need. Firstly, the psychometric qualities of currently available comprehensive norm-referenced spoken language measures were reviewed to provide information to assist SLPs when selecting language measures for use. Secondly, consensus on a taxonomy of terminology for describing child language assessment practices was obtained and strategies to facilitate consistent application of taxonomy terminology were identified. This taxonomy helps to address problems that inconsistent use of terminology poses for the profession. Thirdly, survey data was collected on SLPs current language assessment practices for school-aged children to examine the alignment between current practice and evidence-based practice. In this final chapter, contributions from all the studies in this thesis are summarised and future actions and research directions are identified.

**7.1. Knowledge Gained from this Thesis**

**7.1.1. Research area one: Psychometric properties of child language measures.**
Evidence-based practice recommendations identify that SLPs should use language measures that have the best psychometric evidence when diagnosing school-aged children with language disorder (Betz, Eickhoff, & Sullivan, 2013; Dockrell & Marshall, 2015; Glover & Albers, 2007). Although previous studies have examined the psychometric properties of norm-referenced child language measures (Betz et al., 2013; Friberg, 2010), researchers identified the need for a systematic review that consolidates all available information to make informed recommendations for clinical practice. The study reported in Chapter 2 of this thesis helped to fill this need as the first systematic review to have investigated the psychometric properties of comprehensive norm-referenced spoken language measures for primary school-aged children. This review was the first study of child language measures to examine

independently published studies as well as studies presented in assessment manuals. The inclusion of independent studies allowed for all evidence to be considered when making recommendations for use of language measures (Marshall, Goldbart, Pickstone, & Roulstone, 2015).

Findings from this systematic review in Chapter 2 highlighted the need to improve the design and reporting of studies examining the psychometric properties of comprehensive norm-referenced spoken language measures. Only one quarter of the studies included in the review were identified as having sufficient methodological quality and, due to methodological weaknesses, none of the reviewed language measures presented with evidence of structural validity, internal consistency or measurement error. Furthermore, there was a striking absence of studies employing statistical methods based on item response theory (IRT). IRT methods offer advantages over statistical analyses involving Classical Test Theory (CTT), particularly when establishing appropriate ordering of items in terms of language difficulty (Schmitt, Logan, Tambyraja, Farquharson, & Justice, 2017). Therefore, the absence of IRT methods represents a weakness in the overall quality of studies examining the psychometric properties of language measures (Daub, Skarakis-Doyle, Bagatto, Johnson, & Cardy, 2019). This systematic review also highlighted that the majority of psychometric studies are contained in the manuals produced by publishing companies who have a commercial interest in the language measure. This means that for most measures, independent peer-reviewed psychometric evidence is limited. It also means that SLPs are not able to view studies examining the psychometric evidence prior to purchasing the measure.

Of the comprehensive norm-referenced spoken language measures included in this systematic review, the measures with the best evidence for diagnostic use were the Assessment of Literacy and Language (ALL; Lombardino, Leiberman, & Brown, 2005), Clinical Evaluation of Language Fundamentals-5th Edition (CELF-5; Wiig, Semel, & Secord,

2013), Clinical Evaluation of Language Fundamentals: Preschool – 2nd Edition (CELF:P-2; Wiig, Secord, & Semel, 2004), and Preschool Language Scale -5th Edition (PLS-5; Zimmerman, Steiner, & Pond, 2011). Given the Assessment of Literacy and Language (ALL) only has normative data corresponding to grade levels in the United States of America, the Clinical Evaluation of Language Fundementals-5th Edition (CELF-5), Clinical Evaluation of Language Fundementals: Preschool-2nd Edition (CELF:P-2) and PLS-5 are recommended for diagnostic use in Australia. However, the finding that all the reviewed assessments present with significant limitations in psychometric evidence further emphasises the recommendation that SLPs should collect assessment data from multiple sources when identifying children with language disorder so as to reduce the risk of misdiagnosis.

**7.1.2. Research area two: Terminology for describing types of child language assessments.** Inconsistent use of terminology is identified as a problem across the speech language pathology profession (Walsh & IGOTF-CSD., 2006). Lack of consistent terminology precludes accurate data collection on SLPs current clinical practices and hinders the ability to make detailed comparisons between different assessment practices. Medico-legal responsibilities may also be compromised as clinical assessment practices may not be described in case notes with sufficient clarity for others to accurately interpret (Cameron & Turtle-Song, 2002; Cowie et al., 2001).

The need for well-defined and agreed-upon terms for describing language assessment practices was addressed in the study reported in Chapter 3 of this thesis. In a Delphi study, 55 Australian SLPs reached a high level of consensus on a taxonomy with terminology for describing child language assessment practice. The agreed upon terminology is presented in Supplementary Appendix 3.1. A Delphi study was an ideal methodology for achieving this consensus, as this method allowed for opinions from a large number of participants to be

collected and combined to reach a decision (Boulkedid, Abdoul, Loustau, Sibony, & Alberti, 2011).

In Chapter 4 of this thesis, application of the taxonomy was further explored. Semi-structured interviews were conducted with 13 Delphi study participants and thematic analysis was used to examine participants' perceptions of factors that may influence application of the taxonomy and identify strategies that may facilitate consistent use of the taxonomy by SLPs. The use of semi-structured interviews allowed for more in-depth exploration of SLPs' perspectives regarding the implementation of the taxonomy than was possible in the online Delphi study survey (DiCicco-Bloom & Crabtree, 2006). As no other taxonomy for describing child language assessment practices has agreement from a large group of SLPs, findings from this study represent a significant advancement for the field of child language assessment. The taxonomy developed in this study can be used in situations that require consistent descriptions of assessment practices, for example survey research, quality assurance projects and case notes. The terminology in the taxonomy also has applications in both undergraduate training and continuing professional development, as it may facilitate shared understandings and promote reflection on the different features of different language assessment practices. The strategies that were identified in Chapter 4 provide practical guidance for SLPs who are seeking to establish consistent use of the taxonomy in professional settings.

**7.1.3. Research area three: Profiling current Australian SLP language assessment practice.** Data on current language assessment practice is important for determining the extent to which current clinical practice aligns with evidence-based practice recommendations (Caesar & Kohler, 2009). Previous surveys of SLPs' child language assessment practices have been conducted; however, these surveys have predominantly focused on SLPs in specific agencies, specific populations of children or the use of single

language measures or assessment procedures (refer to Table 1.1. for list of previous surveys). The study reported in Chapters 5 and 6 of this thesis added to current literature by providing data on the range of language assessment practices used by Australian SLPs from different service agencies for a broad population of children. The taxonomy developed in Chapter 3 was used in the survey to facilitate consistent descriptions of assessment practices by survey participants. Furthermore, the strategies for facilitating consistent application of the taxonomy that were identified in Chapter 4 were incorporated into the survey design. As such, this survey underwent more robust development than previous surveys in the field of child language and represents a methodological advancement in survey design.

Strategies to facilitate consistent application of terminology in the survey included pre-categorising a wide range of language measures and assessment procedures using taxonomy terminology for survey participants and only asking about one type of assessment per survey question to reduce cognitive load. Survey participants were also able to complete the survey over more than one sitting if they wished. Definitions for all terms were included with survey questions and focus was placed on the key aspects of definitions rather than taxonomy terms themselves to reduce pre-conceived assumptions by survey participants regarding the meaning of terms. Survey participants were also explicitly instructed to use the definitions provided in the survey, even if they define terms differently themselves.

Results from Part I of the survey were reported in Chapter 5. A total of 407 Australian SLPs provided information on the regularity with which they use different types of assessments, the main challenges they experience concerning language assessment, and primary sources from which they obtain information on language assessment practice. It was identified from the survey that SLPs regularly use assessments that are described as norm-referenced, de-contextualised and conducted in a clinical context and less regularly use assessments that are contextualised, activity-focused, dynamic or conducted in home or

school contexts. Results from regression analysis determined that 'service agency', 'Australian state' and 'years since graduation' may influence the regularity with which SLPs use some types of assessments. Differences were also identified with regards to the main challenges experienced by SLPs in different service agencies when assessing the language abilities of school-aged children.

As no previous survey has examined the influence of different factors on SLPs use of various types of assessment practices, data from this survey contributed new information that shapes greater understanding of the contextual factors that influence SLPs child language assessment practices. In particular, findings from this survey enhance current understanding with regards to the influence of service agency on SLPs assessment practice. A previous study examining the assessment practices of SLPs employed in education agencies identified that workplace policy relating to eligibility criteria is a primary reason for use of norm-referenced language measures by this population of SLPs (Fulcher-Rood, Castilla-Earls, & Higginbotham, 2018). However, the comparison between service agencies reported in Chapter 5 of this thesis identified that SLPs in education agencies do not use norm-referenced language measures more regularly than SLPs in other agencies, such as private practice, that are unlikely to have strict eligibility criteria for service provision. This finding indicates that the factors that influence SLPs' assessment practice are more complex than previous literature has suggested.

Data from this survey also provides greater insight into the role that postgraduate qualifications and SLPs' years of working experience play in influencing speech language pathology practice. Although a previous Australian study identified that procession of postgraduate qualification influences SLPs' engagement in research activities (Finch, Cornwell, Ward, & McPhail, 2013), results from this survey indicated that possession of postgraduate qualifications does not influence the regularity with which SLPs use different

types of assessments. This indicates that higher qualifications do not appear to change SLPs clinical assessment practice.

Years of working experience has not been identified in previous surveys as a factor that influences SLP assessment practice (Caesar & Kohler, 2007; Hux, Morris-Friehe, & Sanger, 1993; Pavelko, Owens, Ireland, & Hahs-Vaughn, 2016; Roulstone et al., 2015). However, the use of regression analysis reported in Chapter 5 of this thesis identified that although SLPs' years of working experience was not significant at a univariate level, this factor significantly contributed to influencing the regularity with which contextualised and dynamic gradual prompting assessments were used when combined with Australian state in multivariate analysis. This identifies that the factors that influence SLPs' language assessment practice are complex and inter-related. These findings also illustrate the need to conduct multivariate analyses to examine the factors that influence assessment practice, as the interaction between complex factors may produce significant findings.

Part II of the survey was reported in Chapter 6. In this part, 335 of the same survey participants provided information on the specific language measures (i.e., measures with set guidelines for administration and scoring), assessment procedures (i.e., procedures without set guidelines for administration and scoring) and assessment delivery methods they use when assessing the language abilities of school-aged children. Data was also collected on the domains assessed, purposes of use and reasons for which language measures are chosen for use. It was identified from the survey that over 130 language measures are used for assessing the language abilities of school-aged children, however, only a small number (approximately seven) are used by each individual SLP. SLPs appear to favour language measures that target semantics and syntax in word and sentence level tasks and thus may be missing important information on children's social abilities and discourse skills. In addition, SLPs may not be

routinely collecting information from parents and teachers when assessing children, which may lead to functional performance in everyday activities not being adequately assessed.

An encouraging finding in this survey was that SLPs reported selecting diagnostic language measures based on psychometric properties; however, this finding was not identified in the case of screening measures. Furthermore, it was identified that SLPs may be using measures to target domains that the measures are not ideally suited to measuring, which may influence the accuracy with which assessment results are interpreted. It was also identified that SLPs use other personnel to conduct assessments, despite limited evidence to support this practice. In contrast, use of ICTs as a method of conducting language assessment was identified as potentially being under-utilised.

## 7.2. Alignment Between Clinical Practice and Evidence-Based Practice

This thesis opened with a quote that reminds us of the importance of continually reflecting upon and improving our clinical practice if we are to continue improving outcomes for the children we serve:

*"Without continual growth and progress, such words as*

*improvement, achievement and success have no meaning"*

– Benjamin Franklin

Assessment practices that are effective in identifying the needs of children are an important component in service provision for children with language disorder. It was a desire to improve language assessment practice for school-aged children that provided the impetus for investigating the objective in this thesis:

*To identify future actions and research directions that may facilitate implementation of*

*evidence-based practice recommendations for child language assessment.*

Overall, findings from this thesis identify that current clinical assessment practice for school-aged children is not well-aligned with evidence-based practice recommendations. Subsequently, a number of recommendations for improving clinical assessment practice are identified. These recommendations are discussed in the following section. Definitions of the terms used in the recommendations to describe assessments are provided in Supplementary Appendix 3.1.

**7.2.1. Using assessments that target different contexts and tasks.** In Chapter 5 it was identified that most SLPs regularly use assessments that are norm-referenced, de-contextualised and conducted in a clinical context and less regularly use other types of assessments. Although norm-referenced language measures provide important data on a child's language abilities, research identifies that data from norm-referenced measures should be supplemented with data from other types of assessments (Bishop & McDonald, 2009; Volden et al., 2017). Therefore, actions are needed to increase the regularity with which SLPs use a range of different types of assessments when assessing the language abilities of school-aged children.

**7.2.2. Selecting language measures to match the intended purpose.** The survey results presented in Chapter 6 identified that SLPs are using language measures for purposes for which the measures may not be designed and validated for. For example, 47-58% of SLPs who reported regularly using the core language subtests from versions of the Clinical Evaluations of Language Fundamentals (CELF) indicated that 'selecting interventions' was a primary purpose for which these measures are used. However, literature identifies that comprehensive norm-referenced language measures such as the Clinical Evaluation of

Language Fundamentals (CELF) are not suited for this purpose. Similarly, over 80% of SLPs who regularly used versions of the Renfrew Action Picture Test (RAPT; Renfrew, 2010) reported using this measure as a screening measure, despite this measure not having evidence of diagnostic accuracy. Furthermore, the core and language index subtests of the Clinical Evaluation of Language Fundamentals (CELF) and the Renfrew Action Picture Test (RAPT) were used to assess 'social-abilities and discourse' by 13-16% of SLPs who regularly use these measures, even though these measures are not designed to target this domain. These survey findings are concerning because assessment data should be collected using appropriate language measures to ensure that the clinical decisions made from the assessment data are as sound as possible. Therefore, it is necessary that actions are taken to assist SLPs with selecting language measures that are appropriate for the domains being measured and the purposes for which assessment data are to be used.

### 7.2.3. Selecting appropriate assessments for children from CALD backgrounds.

In Chapter 5 it was identified that SLPs' assessment practices do not appear to be influenced by the proportion of children on the SLPs caseload from CALD backgrounds. In the survey, SLPs who reported that over half of the last 40 children they assessed were from CALD backgrounds did not report using norm-referenced language measures less regularly, nor did they report using alternatives such as dynamic assessment more regularly. This suggests that SLPs are not altering their assessment practices in response to cultural and linguistic diversity in a child's background. As research identifies that norm-referenced language measures that are normed for monolingual English speaking children are unsuitable for children who do not match this background (Arias & Friberg, 2015; Pearce & Williams, 2013), it is necessary that focus be placed on improving SLPs' assessment practice for children from CALD backgrounds.

### 7.2.4. Collecting data from parents and teachers as an assessment method.

Findings from Chapter 6 indicated that, surprisingly, most SLPs do not report routinely collecting information from parents and teachers when assessing the language abilities of school-aged children. Given the identified limitations of norm-referenced language measures, SLPs may miss diagnostically significant information on a child's language performance if reports from parents and teachers are not also collected during the assessment process (Bishop & McDonald, 2009). Obtaining information from parents and teachers also helps to ensure that SLPs have the information necessary to plan interventions that specifically target a child's functional communication needs in everyday contexts (Tager-Flusberg et al., 2009; Thomas-Stonell, Washington, Oddson, Robertson, & Rosenbaum, 2013; Trembath, Westerveld, & Shellshear, 2016). Furthermore, involving parents in the collection of assessment data may assist with establishing family engagement, which is recognised as being important in maximising service outcomes (Crais, Roy, & Free, 2006). For these reasons, it is necessary that actions are taken to increase the regularity with which SLPs collect data from parents and teachers when assessing the language abilities of children.

### 7.2.5. Using ICTs as an assessment method. 
In Chapter 6 it was reported that only a small proportion of SLPs are using ICTs to deliver language assessment services, despite research indicating that results from language assessments conducted by ICTs are comparable to results from language assessments conducted face-to-face (Waite, Theodoros, Russell, & Cahill, 2010a, 2010b). Problems relating to limited service options and long travel times for paediatric speech pathology services, including services for children with language disorder, have been documented in literature (O'Callaghan, McAllister, & Wilson, 2005; Ruggero, McCabe, Ballard, & Munro, 2012). As use of ICTs provides a means of reducing travel time, increasing SLPs' use of ICTs when conducting assessment services is important for enhancing service accessibility for children and their families (Mashima & Doarn, 2009).

**7.2.6. Using other personnel to conduct assessments.** Another finding reported in Chapter 6 was that SLPs are delegating the role of conducting assessments to other personnel. Nearly one quarter (22%) of SLPs reported that teachers or teacher-aides conduct assessments with at least some of the children on their caseload. This finding raises important questions as to whether teachers or teacher-aides are suitably trained to conduct the assessments they are conducting, and consequently, whether results of these assessments can be considered accurate. Until further knowledge is gained regarding the reliability of having other personnel conduct language assessments, SLPs should be aware that they are responsible for services conducted by teachers or teacher-aides on their behalf, and as such, be cautious in delegating this role (American Psychological Association, 2000; American Speech and Hearing Association, 1997-2018; Speech Pathology Australia, 2016).

## 7.3. Future Actions to Facilitate Implementation of Evidence-Based Practice Recommendations

It is important that evidence-based recommendations for child language assessment practice are implemented by SLPs in clinical practice. However, progress towards this aim appears to be slow. In 1995, a survey by Beck identified an over-reliance on norm-referenced language measures by school-based SLPs in the United States of America, particularly for the purpose of selecting intervention goals. Other authors have also previously identified the need to increase SLPs' use of contextualised assessments when assessing the language abilities of school-aged children (Kemp & Klee, 1997; Pavelko et al., 2016; Trembath et al., 2016; Westerveld & Claessen, 2014) and stop use of norm-referenced language measures for children from CALD backgrounds (Caesar & Kohler, 2007; Teoh, Brebner, & McAllister, 2017). This history suggests that significant barriers exist for SLPs in relation to successful implementation of evidence-based practice recommendations for child language assessment To hasten change in current practice, it is proposed that explicit focus be placed on

identifying and addressing these barriers (Hakkennes & Dodd, 2008; Olswang & Prelock, 2015).

In Chapter 1 of this thesis a framework outlining factors that may influence implementation of evidence-based practice recommendations was introduced. This framework is presented again in Figure 7.1.
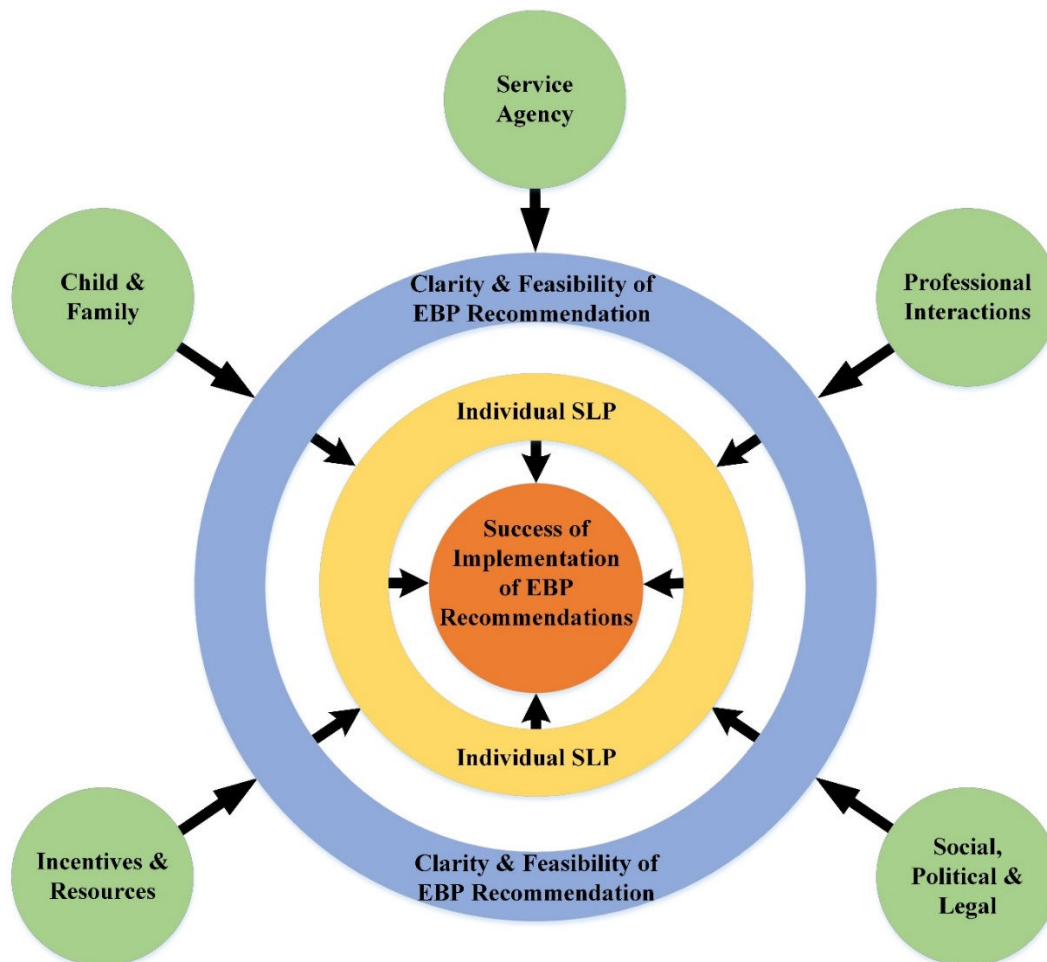


*Figure 7.1.* Factors that may influence SLPs implementation of evidence-based practice recommendations (based on the TICD checklist by Flottorp et al., 2013).

The factors presented in the figure are based on the comprehensive, integrated checklist of determinants of practice (the TICD checklist) which was developed to assist with identifying barriers and facilitators that may influence health practitioners' implementation of evidence-based practices (Flottorp et al., 2013). In the following section, the factors in Figure

7.1 are reflected upon in relation to new knowledge obtained through this thesis. Using this framework, specific actions to improve implementation of evidence-based practice recommendations for child language assessment are identified and discussed.

**7.3.1. Factors related to the individual SLP.** Implementation of evidence-based practice recommendations may be influenced by factors related to the individual SLP as shown in the inner yellow ring of Figure 7.2.
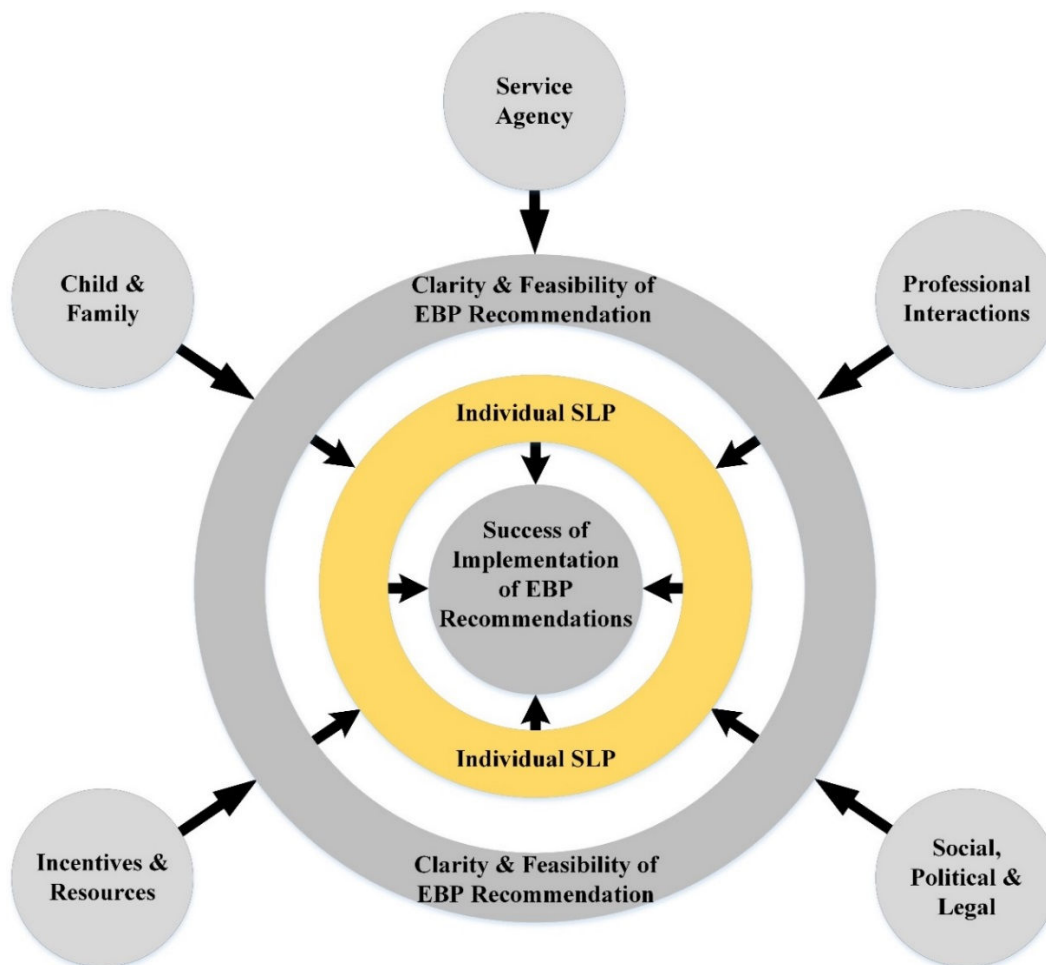


*Figure 7.2.* Factors internal to individual SLPs that may influence implementation of evidence-based practice recommendations (based on the TICD checklist by Flottorp et al., 2013).

In Chapter 5 of this thesis it was identified that SLPs do not frequently use journal articles as a source of information on language assessment practice, but instead rely on information from peers. This finding has also been identified earlier surveys examining SLPs

assessment practice (Beck, 1995; Wilson, Blackmon, Hall, & Elcholtz, 1991) and studies examining SLPs use of evidence-based practices more generally (Hoffman, Ireland, Hall-Mills, & Flynn, 2013; O'Connor & Pettigrew, 2009; Vallino-Napoli, 2004). Given that journal articles are likely the most up-to-date sources of information on clinical practice, low utilisation of journal articles may lead to SLPs having a lack of awareness of evidence-based practice recommendations for language assessment and the evidence that supports the recommendations.

One reason for low use of journal articles may be limited access to journal articles in workplaces (Cheung, Trembath, Arciuli, & Togher, 2013). As such, there is an identified need for both individual SLPs and professional associations to advocate for actions that improve SLPs access to, and active use of, journal articles (Cheung et al., 2013; Harding, 2014; O'Connor & Pettigrew, 2009). To assist with increasing awareness of evidence-based practice recommendations, research knowledge should also be disseminated though other avenues (Laver, Brown, Cordier, & Lannin, 2018). In Chapter 5, workshops or formal presentations were identified as the second most frequently reported sources of information on assessment practice and may, therefore, be an effective means of disseminating research evidence to SLPs. It is also important that professional speech language pathology associations take an active role in ensuring that information presented in professional development events is linked to recently published research studies, for example, by subjecting the content of workshops to peer-review (Van Achterberg et al., 2006).

Although SLP attitudes and beliefs were not specifically targeted in this survey, findings reported in Chapter 6 of this thesis identified that, although over 130 language measures are in use by SLPs as a group, each individual SLP uses on average only about seven different assessments. One reason for this finding may be that SLPs have personal preference for a specific language measure, even though this measure may not be ideally

suited for all purposes. For example, SLPs may always use a familiar language measure because it is easier to always use the same measure or SLPs may continue to use a measure that was expensive to purchase to return the cost of investing in the measure (McCabe, 2018). To facilitate implementation of evidence-based assessment practices, SLPs should be supported during under-graduate training and continuing professional development to engage in critical reflection regarding their own assessment practice and the influence of their own beliefs and attitudes on their assessment choices (Michie et al., 2005; Perkins et al., 2007; Shrubsole, Worrall, Power, & O'Connor, 2018).

Implementation of evidenced-based practice recommendations may be influenced by the knowledge and skills of SLPs (Michie et al., 2005). In Chapter 6 this thesis, lack of skills and confidence was reported by SLPs as a challenge when conducting language assessment, particularly when conducting assessments for children from CALD backgrounds. This theme also emerged in thesis Chapter 4, with some semi-structured interview participants commenting that dynamic assessment is not well understood by SLPs generally across the profession. Furthermore, previous surveys of SLP child language assessment practice have also identified that lack of familiarity and training in dynamic assessment poses a barrier to evidence-based assessment practice for children from CALD backgrounds (Arias & Friberg, 2015). Given these findings, professional development in conducting different types of assessments, particularly dynamic assessment, should be a priority for SLPs in both under-graduate training and in continuing professional development (Teoh et al., 2017). As findings from Chapter 6 identified that SLPs may be using language measures for purposes for which the measures are not suitable, professional development for SLPs should also include training in selecting language measures to match the purpose for which the measure will be used.

When providing professional development to SLPs, consideration needs to be given to how professional development is provided. It is widely accepted that one-off training sessions

or provision of written materials alone are unlikely to result in effective implementation of practice recommendations (Burke & Hutchins, 2008; Ratner, 2006; Wandersman et al., 2008). Instead, professional development needs to include coaching, which involves the provision of context-specific support over an extended period to support practice change (Hoffman et al., 2013; Wandersman et al., 2008). Key components in coaching include facilitation of self-reflection, demonstration of new assessments in practice context and non-evaluative feedback (Ladyshewsky, 2010).

### 7.3.2. Factors related to clarity and feasibility of practice recommendation.

Implementation of practice recommendations may be influenced by factors related to the clarity and feasibility of practice recommendations themselves. This is as shown in the outer blue ring in Figure 7.3.

The systematic review conducted in Chapter 2 of this thesis identified that limitations exist with regards to the psychometric evidence that is available for comprehensive norm-referenced spoken language measures for school-aged children. The majority of studies included in the review were found to have specific weaknesses in methodological quality which compromised the outcomes of the studies. The systematic review also revealed a problem of psychometric evidence being almost exclusively confined to studies produced by assessment publishers who have a commercial interest in the measure. In addition, studies were predominantly published in assessment manuals, rather than peer reviewed journal articles which exposes the psychometric reporting to high risk of bias.
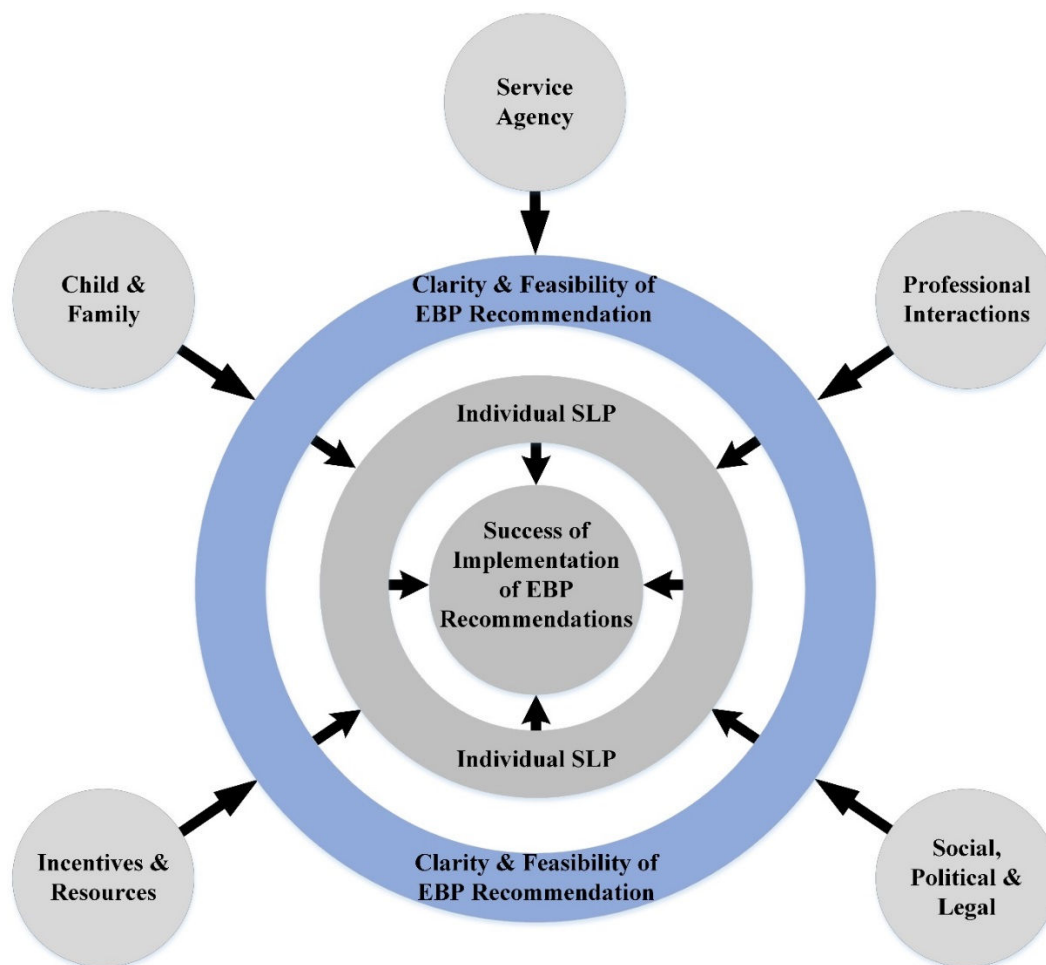
*Figure 7.3.* Clarity and feasibility of the recommendations themselves as factors that may influence SLPs implementation of evidence-based practice recommendations (based on the TICD checklist by Flottorp et al., 2013).

To make sound recommendations for evidence-based practice, research studies with high methodological rigour are needed. Researchers and assessment developers are strongly urged to use guidelines that have international consensus on terminology and methodology, such as the COSMIN guidelines, when designing and reporting on psychometric studies (Mokkink, Prinsen, Bouter, De Vet, & Terwee, 2015) and subject psychometric studies to independent peer review before publication. To drive change in quality of psychometric evidence for language measures, SLPs and professional associations need to advocate for improved practices in creation and dissemination of psychometric evidence (Daub et al., 2019). For example, actively insisting that assessment developers improve the

methodological standards of psychometric studies and contacting publishers to request information on psychometric evidence before purchasing a language measure.

Even when SLPs have access to research evidence, they may lack the knowledge needed to synthesise research evidence for application into their clinical context (Finch et al., 2013; Harding, 2014; Metcalfe et al., 2001). One reason for this may be the time and effort needed to read and analyse information from multiple research studies (Hoffman et al., 2013). To reduce time and effort, it is important that SLPs are provided with literature that summarises all available information pertinent to a practice recommendation (Flottorp et al., 2013; Nail-Chiwetalu & Ratner, 2006). Inconsistencies between different sources of information also need to be identified and explained to assist SLPs in making evidence-based decisions. For example, a published review of the Clinical Evaluation of Language Fundamentals-5th Edition (CELF-5) cautioned SLPs against using of the Clinical Evaluation of Language Fundamentals-5th Edition (CELF-5) due to psychometric limitations (LEADERS, 2014). This review was limited to an evaluation of the psychometric studies reported in the assessment manual, did not use a framework for evaluating the methodological quality of studies, and did not make comparisons between the Clinical Evaluation of Language Fundamentals-5th Edition (CELF-5) and other language measures. The recommendation from this review may appear to be contradictory to the recommendation arising from the systematic review reported in Chapter 2 of this thesis, however, although the Clinical Evaluation of Language Fundamentals-5th Edition (CELF-5) has psychometric limitations, it presents as one of the diagnostic measures that have the best evidence for use when systematically compared to other measures.

To support SLPs in synthesising research evidence for application into clinical practice, further systematic reviews are needed to comprehensively gather and summarise all available research findings (Marshall et al., 2015; Nail-Chiwetalu & Ratner, 2006).

Resources that assist SLPs in sourcing and evaluating evidence-based information are also needed, for example, resources similar to the SpeechBite website, which rates intervention studies according to methodological quality (http://speechbite.com). Furthermore, the development of clinical practice guidelines are needed to explicitly outline how research findings are translated into everyday workplace actions and processes (Hoffman et al., 2013; Ratner, 2006).

It is also important that practical barriers to implementing recommendations in everyday clinical practice are addressed within evidence-based practice recommendations (Green, Ottoson, Garcia, & Hiatt, 2009). For example, previous surveys have identified that SLPs experience barriers to implementing evidence-based practice recommendations for children from CALD backgrounds due to the non-existence of required resources, such as suitable interpreters, bilingual SLPs or developmental norms that are appropriate to the child's home language (Arias & Friberg, 2015; Guiberson & Atkins, 2012; Kadyamusuma, 2016; Teoh et al., 2017). For practice recommendations to be feasible to implement, it is important that they are accompanied by specific recommendations as to how to over-come barriers that may be present in day-to-day clinical practice of SLPs. Pilot studies examining practical implementation of evidence-based practice recommendations may be helpful in identifying, and subsequently addressing, practical barriers related to implementation.

**7.3.3. Factors external to individual SLPs that influence implementation.** Factors external to the SLP are related to service agency, child and family, professional interactions, resources and incentives or social, political and legal issues. These are depicted in the outer green circles in Figure 7.4.
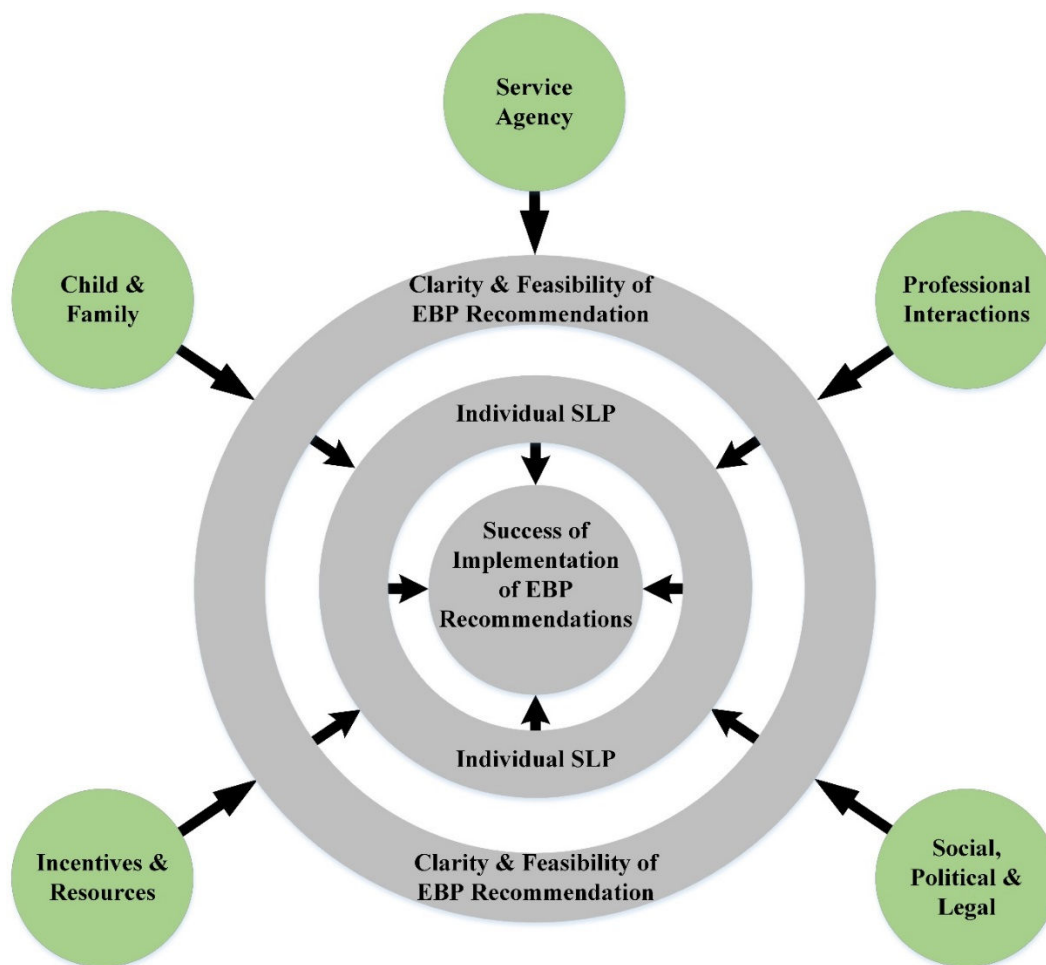
*Figure 7.4.* Factors external to individual SLPs that may influence SLPs implementation of evidence-based practice recommendations (based on the TICD checklist by Flottorp et al., 2013).

**7.3.3.1. Service agency.** The findings reported in Chapter 5 of this thesis indicated that service agency is a factor that significantly influences the regularity with which SLPs use some types of language assessments. Significant differences were also identified in relation to the frequency with which SLPs in different agencies reported particular challenges as main challenges when assessing the language abilities of school-aged children. For example, 'limited time to meet with teachers' was more likely to be reported as a main challenge by SLPs in education agencies compared to SLPs in other agencies and 'setting constraints' (i.e., not able to see children in particular locations) was more likely to be reported by SLPs in health agencies compared to other agencies. These findings highlight that barriers and

facilitators to implementation of evidence-based practice recommendations may be different across service agencies. When identifying actions to improve implementation of practice recommendations, it is important that SLPs and service managers assess the unique challenges that exist within their specific service agencies (Hakkennes & Dodd, 2008; Nail-Chiwetalu & Ratner, 2006).

Workplace culture and leadership may also influence uptake of evidence-based practices (Belkhodja, Amara, Landry, & Ouimet, 2007). In Chapter 5 of this thesis, 'information from employer' was the third most frequently reported main source from which information on assessment practice was obtained. This finding highlights the important role that service managers can play in facilitating implementation of evidence-based practice recommendations. However, in previous literature, health professionals have reported that evidence-based practices are not always supported by managers, with service effectiveness often measured purely by number of service-users seen rather than the quality of the service provided (Harding, 2014). To support implementation of evidence-based practice recommendations, it is proposed that service agencies actively seek to develop a workplace culture that encourages evidence-based practice (Rappolt & Tassone, 2002). In addition to ensuring that service managers are well-acquainted with evidence-based practice recommendations, it is important to ensure that service managers have the leadership skills needed to instigate practice change (Flottorp et al., 2013). This includes providing managers with explicit training in implementation science to build capacity within workplaces. For example, training on how to instigate a process of change by identifying areas in need of change, designing strategies to address need for change, clearly defining roles and performance expectations and monitoring progress in achieving change (Graham et al., 2006).

*7.3.3.2. Child and family*. A finding from Chapter 5 of this thesis was that SLPs did not frequently report child and family expectations regarding service provision as a main challenge when assessing the language abilities of children. This suggests that SLPs perceive this factor as less challenging compared with other factors. However, although not identified as a main challenge in this survey, expectations from children and families may still be an influencing factor. For example, previous studies have identified that families report problems with unacceptably long waiting lists and unaffordable service costs in relation to paediatric SLP services (O'Callaghan et al., 2005; Ruggero et al., 2012). Given that findings in this survey identified issues related to 'limited time' as frequently reported main challenges, it is possible that SLPs respond to pressure from families by prioritising time and cost efficiency over other reasons when selecting language measures or assessment procedures for use (Harding, 2014). Furthermore, an earlier study identified that expectations of families may vary depending on whether children access services through health or education agencies, with parents having greater expectation for collaboration between SLPs and teachers if services were provided through an education agency (Carroll, 2010). Therefore, it is possible that differences between service agencies with regards to assessment practice are influenced by different child and family expectations.

A notable finding from Chapter 6 was that most SLPs do not regularly collect data from parents when assessing the language abilities of school-aged children, suggesting that SLPs are not collaborating with families to an ideal extent during the assessment process. A main goal of assessment should be to establish a shared understanding with families regarding a child's abilities, needs and service requirements (Dunst, Leet, & Trivette, 1988). Such a goal is unlikely to be achieved when collaboration with families is limited. Furthermore, by not collecting information from families, SLPs may make erroneous assumptions about the expectations that families have in relation to service provision. In a

previous study examining the perceptions of SLPs and service users regarding ICTs as a method of service provision, it was identified that service-users had much greater access to technology and more positive attitudes towards ICTs than SLPs assumed service-users would have (Dunkley, Pattie, Wilson, & McAllister, 2010). Similarly, another study comparing the perceptions of SLPs and children with language disorder regarding ideal speech language pathology service provision identified that, while most SLPs focused services on improving child's language skills, most children valued services that directly targeted participation and inclusion at school (Gallagher, Murphy, Conway, & Perry, 2019).

To ensure that the service expectations between SLPs, children and families are aligned, there is an identified need for SLPs to engage in family-centred practices when assessing the language abilities of school-aged children (Crais et al., 2006; Gillon et al., 2017). To do this, SLPs may need to develop skills in 'actively listening' to children and their families during the assessment process and incorporating this information into service planning and clinical decision-making (Gallagher et al., 2019). This includes specific skills in methods for interviewing children (Owen, Hayett, & Roulstone, 2004).

*7.3.3.3. Professional interactions.* In Chapter 5 it was identified that most SLPs reported 'informal discussions with peers and colleagues' as one of their primary sources of information on assessment practice. This finding is similar to previous studies, which have also identified that SLPs use information from peers as a primary source of information on clinical practice (Beck, 1995; O'Connor & Pettigrew, 2009; Wilson et al., 1991). Given this trend, it is important that actions are taken to facilitate sharing of evidence-based information (Nail-Chiwetalu & Ratner, 2006). This includes establishing processes in workplaces for sharing research articles and making sense of the research findings, for example, by establishing journal clubs (Hoffman et al., 2013; Nail-Chiwetalu & Ratner, 2006). It is also important that undergraduate training and continuing professional development is focussed

on developing SLPs' skills to critically appraise information and, as such, seek evidence to back up assessment practices that are recommended by peers (Hoffman et al., 2013; Nail-Chiwetalu & Ratner, 2006).

Individual SLPs may be more likely to change practice if those around them are making the same changes (McCabe, 2018). In a study by Harding (2014), health practitioners felt more able to address practice change when working as part of a group. Working as a group also creates greater accountability in terms of committing to changes that the group has agreed-upon (Flottorp et al., 2013). Therefore, actions to support the implementation of practice recommendations for language assessment include establishing peer support networks to facilitate shared learning, for example learning to conduct a new language measure or assessment procedure at the same time as a peer (McCabe, 2018).

*7.3.3.4. Resources and incentives.* Findings from Chapter 5 identified that 'limited time' and 'limited assessment materials' (e.g. due to budget constraints) are main challenges that SLPs frequently report when conducting language assessment. Consistent with 'limited time', being 'quick to administer and score' also appeared as a main reason in Chapter 6 for selecting screening measures. Furthermore, these same challenges are frequently reported in previous surveys of SLPs assessment practice (Arias & Friberg, 2015; Fulcher-Rood et al., 2018; Guiberson & Atkins, 2012; Huang, Hopkins, & Nippold, 1997; Pavelko et al., 2016; Westerveld & Claessen, 2014). For this reason, specific attention needs to be given towards identifying actions that address 'limited assessment materials' and 'limited time'. This includes further investigations to identify the specific aspects of assessment practice that are time-consuming. It is widely recognised that more effort is required to learn a new practice compared with continuing with an old practice (McCabe, 2018; Samuelson & Zeckhauser, 1998). Therefore, it is possible that the challenges reported by SLPs in relation to 'limited time' are associated with learning to conduct a new language measure or assessment

procedure, rather than administering the measure or procedure itself; however, further investigation is needed to substantiate this notion.

It is also possible that some language measures or assessment procedures can be conducted in a more time efficient way than the manner in which SLPs currently conduct these. For example, research has identified that the time required to conduct contextualised assessments of social abilities and discourse can be reduced by collecting shorter language samples than SLPs may typically collect in clinical practice (Heilmann, Miller, & Nockerts, 2010; Pavelko et al., 2016). Similarly, it is possible that what SLPs perceive as 'limited assessment materials' reflects a lack of SLPs' knowledge in how to use materials that are readily available in a time-efficient manner, although, again, further investigation is needed.

It is important that the time and resourcing needs involved in conducting evidence-based language assessment are explicitly considered and addressed by both individual SLPs and service managers (Harding, 2014). SLPs' caseloads need to allow sufficient time for conducting a quality assessment with each child. This includes time explicitly allocated to professional development activities, such as setting aside time to learn to conduct new language measures or assessment procedures (Arias & Friberg, 2015; Cheung et al., 2013). Making changes to practice through a series of small steps over time may also reduce the need for a high time-commitment in the initial stages (Flottorp et al., 2013).

Findings from Chapter 6 of this thesis identified that many options exist for SLPs when choosing language measures or assessment procedures for use. Choosing from a broad array of possible assessment options, while balancing competing needs such as research evidence, service agency policy, child and family preferences, and time constraints; may be understandably taxing for SLPs. As a consequence, SLPs may experience choice overload, which describes difficulties weighing up all considerations to make a sound decision (McCabe, 2018; Ratner, 2006). To facilitate implementation of evidence-based assessment

practices, SLPs may benefit from decision-support aids, for example a flowchart outlining the steps involved when choosing language measures or assessment procedures for use (McCabe, 2018). The taxonomy developed in this thesis may assist in structuring such a flowchart by providing a structure for reflecting on different assessment practices. Furthermore, service agencies are encouraged to embed decision-support aids into clinical processes, for example, reminders on referral forms or case note entry forms regarding assessment data that needs to be collected (Kawamoto, Houlihan, Balas, & Lobach, 2005).

The incentives that exist for practice change also need to be considered (Flottorp et al., 2013). Health professionals have reported that implementation of evidence-based practice recommendations is not always linked to key performance indicators or associated with career development, thus providing little incentive towards investing effort, and financial resources, in making practice change (Harding, 2014). To facilitate implementation of evidence-based assessment practice recommendations, it is important that service agencies and professional associations provide incentives, both financial and non-financial, that favour uptake of evidence-based practice recommendations (Michie et al., 2005).

*7.3.3.5. Social, political and legal.* In Australia, health and education services fall under the jurisdiction of individual states and considerable differences are documented in relation to speech language pathology services across states (Speech Pathology Australia, 2014). Similarly, literature from the United States of America has also documented differences across states in relation to language assessment requirements for determining service eligibility (Spaulding, 2012). It is possible that the differences between Australian States identified in Chapter 5 of this thesis are related to differences in state policies; however, further investigation is needed to explore this.

Federal funding policies may also influence assessment practice, particularly use of ICTs for delivering assessment services. In Australia, some speech language pathology

services do not attract public or private health rebates if services are delivered via ICTs, rather than face-to-face, which may pose a significant barrier for use of ICTs as an assessment delivery method (Allied Health Professions Australia, 2017). Furthermore, this issue may not be unique to Australia, as similar barriers related to lack of reimbursement for services provided by ICTs have been reported in international literature (Mashima & Doarn, 2009). To support evidence-based assessment practice, it is important that close links are established between researchers, clinicians and policy-makers to ensure that service policies are continually updated to reflect evidence-based practice recommendations (Flottorp et al., 2013). Professional associations should have a role in consulting with policy-makers and advocating for evidence-based practice recommendations to be enacted at a policy-level.

Ethical and legal concerns may also influence assessment practice. For example, concerns regarding confidentiality have been reported as barriers for use of cloud-based ICTs to deliver services (Mashima & Doarn, 2009). Confidentiality may also be a concern with regards to observing children in classrooms or other daily activities when other children are also present. Furthermore, it is possible that SLPs perceive a higher risk of malpractice claims if they provide services that could be seen as *different* to services that have traditionally been provided, regardless of available evidence (Flottorp et al., 2013). Actions to support the implementation of practice guidelines should include the development of technical standards or position papers, for example, guidelines on ethical service provision using ICTs (Flottorp et al., 2013; Mashima & Doarn, 2009). It is also important to ensure that SLPs are provided with accurate information on legislative requirements and risks of mal-practice as part of their continuing professional development (Flottorp et al., 2013).

## 7.4. Adopting an Implementation Science Process

Findings from Chapters 5 and 6 of this thesis highlight that SLP assessment practice is influenced by complex and interacting factors. As such, it is likely that changing SLPs

current language assessment practice to better align with evidence-based recommendations will involve considered actions from professional bodies, researchers, service managers and individual SLPs. Tools such as the TICD checklist should be used to reflect on factors that may influence implementation of evidence-based practice recommendations by SLPs (Flottorp et al., 2013). The factors in the TICD checklist are represented diagrammatically in Figure 1.1 and again in Figure 7.1. These factors can be categorised against the following groupings: the individual SLP; the practice recommendation itself; service agency; child and family; professional interactions; resources and incentives; and social, political and legal factors. These same factors are also summarised in Table 7.1 with examples of reflective questions relating to the influence of each factor on SLP assessment practice and recommended actions for each factor. By explicitly identifying the barriers and facilitators that exist with regards to implementation of evidence-based practice recommendations, actions that may be most needed to induce practice change can be identified and enacted (Flottorp et al., 2013). Implementation science studies should then be undertaken to assess the outcome of implementing actions specifically designed to effect practice change (Hakkennes & Dodd, 2008). Conducting this implementation science process of identifying barriers, enacting actions to change practice and measuring practice change is necessary if the speech language pathology profession is to be successful in improving child language assessment practice into the future.

Table 7.1

*Examples of Considerations for Each Factor and Reflective Questions Relating to Barriers or*

*Facilitators to Implementation of Evidence-Based Practice Recommendations (based on the*

*TICD checklist by Flottorp et al., 2013).*

| Influencing factor | Examples of considerations within each influencing factor | Examples of reflective questions related to each influencing factor | Examples of actions that may address barriers related to influencing factor |
|---|---|---|---|
| Individual SLP factors | - SLP awareness of and familiarity with recommendation<br>- SLP attitudes and beliefs towards recommendation<br>- SLP level of awareness of own practice in relation to EBP recommendation<br>- SLP knowledge and skill<br>- SLP self-efficacy | - Are SLPs aware of EBP recommendations and the evidence that supports them?<br>- Do SLPs believe that implementing the EBP recommendation is a priority and will improve clinical outcomes?<br>- Are SLPs aware of their own practice in relation to EBP recommendations?<br>- Do SLPs have the knowledge and skills required to implement the EBP recommendation?<br>- Are SLPs confident in their ability to successfully change their practice? | - Actions to increase SLP knowledge of EBP recommendations and the reasons for recommendations (e.g., during peer-reviewed conference or workshop presentations)<br>- Actions that target critical reflection by SLP regarding personal attitudes towards assessment practice (e.g., during professional development events or professional supervision)<br>- Actions that develop SLP skills and confidence in selecting and conducting different types of assessments (e.g., coaching in workplaces) |
| Clarity and feasibility of recommendation | - Quality and credibility of evidence behind recommendation<br>- Clarity and accessibility of recommendation<br>- Consistency of recommendation with other recommendations<br>- Application of recommendation to clinical practice<br>- Feasibility of recommendation | - Is evidence for the EBP recommendation sound?<br>- Is the evidence for the EBP recommendation and consequence of not adhering to recommendation explained clearly?<br>- Could conflict between different sources of information be perceived by SLPs in relation to EBP recommendations?<br>- Do SLPs have clarity in relation to the clinical application of the EBP recommendation in their clinical setting?<br>- Are practical suggestions supplied to assist with overcoming barriers to implementation of EBP recommendation? | - Actions that enhance the quality of studies examining psychometric evidence for language measures (e.g., use of methodological guidelines when designing and reporting on studies)<br>- Actions that increase SLPs access to research studies (e.g., open-access publishing)<br>- Actions that assist with summarising and rating the quality of evidence from different sources (e.g., systematic reviews)<br>- Actions that assist in making EBP recommendations feasible in everyday clinical settings (e.g., clinical practice guidelines with specific examples of how practice recommendations can be applied)<br>- Actions that include piloting of EBP recommendations to test feasibility |
| Service agency | - Service agency regulations or policies<br>- Priority given to practice change<br>- Capability of leadership<br>- Authority and accountability structures<br>- Availability of monitoring and feedback on progress of change | - Do workplace policies align with, and support, implementation of EBP recommendations?<br>- Is practice change placed as a priority within service agencies?<br>- Do leaders and managers have the necessary knowledge, influence and leadership style to effect practice change? | - Actions that establish implementation of evidence-based practice recommendations as a priority in workplaces (e.g., setting explicit expectations for implementation of EBP recommendations)<br>- Actions to develop the leadership skills of service managers (e.g., provision of training in implementation science strategies) |

| Influencing factor | Examples of considerations within each influencing factor | Examples of reflective questions related to each influencing factor | Examples of actions that may address barriers related to influencing factor |
|---|---|---|---|
| | - Availability of necessary supports from external agencies | - Are roles and tasks well-defined (i.e., is it clear who is accountable for ensuring implementation of EBP recommendations and how success of implementation is measured)?<br>- Do service agencies have access to external supports necessary to effect practice change? | |
| Child and family | - SLP perceptions of the expectations, preferences and motivation of children and families<br>- Actual expectations, preferences and motivation of children and families | - Do the expectations and preferences of children and families align with EBP recommendations?<br>- Do SLPs perceive that the expectations and preferences of children and families will not align with EBP recommendations? | - Actions that establish shared understandings between SLPs, children and families (e.g., building SLP capacity to provide family centred assessment practices) |
| Professional interactions | - Influence of professional communications<br>- Effectiveness of communication across different levels of service access e.g. between those who take referrals and those who provide services<br>- Knowledge and skills of work teams | - Are positive attitudes to practice recommendations present across peer networks (e.g. during interactions with colleagues)?<br>- Are work teams able to work together to support practice change? | - Actions that establish sharing of evidence-based information (e.g., journal clubs)<br>- Actions that facilitate a team approach to practice change |
| Incentives and resources | - Availability of required time, materials or technology<br>- Availability of professional support to implement EBP recommendation<br>- Incentives and disincentives (financial or non-financial) | - Do SLPs have the required time, materials and technology access to implement EBP recommendations?<br>- Do SLPs have access to the assistance they need to implement EBP recommendations?<br>- What incentives are available to encourage practice change? | - Actions that ensure sufficient time for implementation of EBP recommendations (e.g., explicit allocation of time in SLP workloads for tasks associated with implementation of EBP recommendations)<br>- Actions that reduce the initial time and resourcing needed to implement EBP recommendations (e.g., making practice changes in small steps)<br>- Actions that provide assistance with clinical decision-making involved in implementation of practice recommendation (e.g., decision-making aids, reminders on case note forms or clinical supervision)<br>- Actions that increase incentives for implementation of EBP recommendations (e.g., measuring implementation of EBP recommendations in key performance indicators) |

| Influencing factor | Examples of considerations within each influencing factor | Examples of reflective questions related to each influencing factor | Examples of actions that may address barriers related to influencing factor |
|---|---|---|---|
| Social, political and legal factors | - Funding policies<br>- Opinions of influential people (outside of service agency)<br>- Ethical issues<br>- Liability issues | - Do economic constraints or funding policies influence implementation of practice recommendations?<br>- Are influential people (e.g., politicians and policy-makers) aware of and supportive of EBP recommendations?<br>- Do real or perceived ethical or legal concerns (such as risks of malpractice complaints) influence implementation of practice recommendations? | - Actions to ensure that service policies reflect evidence-based practice recommendations (e.g., input from multiple stakeholders when developing policy)<br>- Actions that provide specific guidance on ethical and legal issues associated with practice recommendations (e.g., development of standards for maintaining confidentiality) |

*Note*. EBP: Evidence-Based Practice

## 7.5. Strengths and Limitations of the Studies in this Thesis

In this section of the thesis, the strengths and limitations of each study are outlined. Limitations for each study are also listed in the discussion sections of Chapters 2-6. In Chapter 2, a systematic review was conducted to investigate the psychometric properties of comprehensive norm-referenced spoken language measures for children. A strength of this review was the use of the COSMIN taxonomy to rate the methodological quality of the included studies. A further strength of this review was the inclusion of studies from both manuals and independent journal articles. A limitation of this review was that it was limited in scope as the psychometric property of responsiveness was not reviewed. As responsiveness refers to the ability of a measure to detect change in status across time accurately, this review does not provide information on the usefulness of comprehensive norm-referenced spoken language measures for purposes of measuring intervention outcomes (Polit, 2015).

In Chapter 3, consensus on a taxonomy with terminology for describing language assessment practices was developed through a Delphi study. A strength of this Delphi study was that qualitative data was examined addition to quantitative data. This allowed for SLP opinions and perceptions to be elicited and incorporated into the development of the taxonomy through consensus. In addition, agreement with the structure and definitions of the

taxonomy was confirmed over two Delphi study rounds, thus giving weight to the agreement. Another strength was the inclusion of case studies, which allowed practical application of the taxonomy to be examined.

Participant drop-out was somewhat of a limitation of the Delphi study. The response rate for each round was between 71.4% and 78.2%, which is above the minimum level of 70% identified in literature as being needed to maintain rigour across a Delphi study (Sumsion, 1998). Nonetheless, this response rate presents as a limitation as it is not possible to determine if dropout was at random. Given that agreement with the taxonomy was established in round one before dropout occurred, this limitation relates specifically to the development of consensus on the categorisation of assessment practices in case studies.

Another limitation is that completion of the Delphi study required a large amount of reading by Delphi study participants, thus placing substantial time and cognitive demand on participants. Although all attempts were made to reduce the cognitive demands, it is possible that the categorisation of case studies was influenced by participant fatigue. In addition, it was not possible to examine the application of all taxonomy categories using two case studies, therefore it is possible that different results may have been obtained if different case studies with different types of assessment practices were used.

In the study reported in Chapter 4, the application of the taxonomy was further explored through semi-structured interviews with Delphi study participants. The inclusion of this study added strength to the survey methodology reported in Chapters 5 and 6 as strategies to support application of the taxonomy were identified before the taxonomy was used in the survey. A limitation with the semi-structured interview study was the demographics of the participant group. As SLPs with experience applying the taxonomy were required, participants in the semi-structured interviews were SLPs who previously completed the Delphi study. However, it is acknowledged that these SLPs might not be representative of

the broader population of SLPs, due to having been recruited for their particular expertise in the field of child language. It is also possible that Delphi study participants who were also willing to participate in a further interview may have had particular perceptions towards the taxonomy, which could have influenced the study outcomes.

Another limitation of the semi-structured interviews was that, due to the study being a part of the PhD thesis, the PhD researcher conducted both the interviews and data analysis. This opens the possibility of bias from the researcher's prior knowledge and experience. This bias was reduced, but not eliminated, through the use of a reflective journal by the PhD researcher and regular discussions with the research team members regarding data interpretation (Thomas & Magilvy, 2011).

In Chapters 5 and 6, a survey was conducted to examine SLPs' language assessment practice for school-aged children. A strength of the survey was the use of the taxonomy developed in Chapter 3 to facilitate consistent descriptions of assessment practices by survey participants. The use of a defined frequency rating scale with numeric qualifiers also helped to ensure consistency in frequency ratings provided by participants (Blais & Grondin, 2011). A further strength of this survey is that it included a large sample of SLPs across Australia, thus increasing the likelihood that survey findings are generalisable to the broader population of Australian SLPs who work with school-aged children.

A limitation of this survey is that, although the total survey sample size was large; groups with smaller overall populations, such as SLPs in Tasmania, Australian Capital Territory and Northern Territory and SLPs working in Universities, had small sample sizes. This limits the interpretation of findings from these smaller groups. In addition, due to small sample sizes, SLPs from Tasmania, Australian Capital Territory and Northern Territory had to be removed from the regression analysis to improve sensitivity of the analysis. Therefore, it was not possible to compare assessment practices in these smaller states against larger

states. It is also acknowledged that this survey consisted only of Australian SLPs which means that findings may not represent assessment practice in other countries.

As with all survey methodologies, findings from this survey rely on participants accurately recalling and reporting the assessment practices they used. This presents as a limitation in Chapter 6 (Survey Part II) which required participants to list the language measures they used. As 2.2% of measures listed by participants were not identifiable due to use of ambiguous acronyms or names of measures, these responses could not be included.

Finally, it is acknowledged that this survey examined SLPs' assessment practice broadly, rather than examining specific aspects of assessment practice in detail. Considering the paucity of previous research related to child language assessment practice, a broad approach to the inclusion of survey questions was selected to obtain a 'big picture' understanding of assessment practice and identify key areas in need of further specific investigation. Therefore, it is acknowledged that this survey has limitations with regards to the breadth with which some aspects of assessment practice were explored.

## 7.6. Further Research Directions

Further research is needed to investigate a number of areas related to language assessment for school-aged children. In relation to psychometric evidence for use of language measures, there is a need for further studies investigating child language measures for evidence of structural validity, internal consistency and measurement error. In particular, research is needed to examine the psychometric properties of language measures using statistical methods based on item response theory (IRT). The psychometric property of responsiveness should also be investigated to determine the suitability of using comprehensive norm-referenced spoken language measures as outcome measures. Additional systematic reviews are needed to examine the psychometric quality of language measures that were outside the scope of this review and there is a need for the development of new

language measures with robust psychometric qualities, particularly measures that are described as contextualised, activity-focused and dynamic.

In relation to the need for use of consistent terminology for describing language assessment practices, further research is needed to extend upon the taxonomy that was developed in this thesis. This includes further studies examining the applicability of the taxonomy to other countries and studies examining strategies that support consistent application of the taxonomy by different groups of SLPs to those sampled in this study, for example, SLP students and new graduates. Studies are also needed to explicitly examine the degree to which the strategies identified by participants in Chapter 4 enhance the consistency with which SLPs apply the taxonomy.

To further develop an understanding of current SLP assessment practice, further research is needed. As service agency and Australian state were found to influence the regularity with which SLPs use some types of assessments, further investigation is needed to more fully investigate the reasons why these factors influence assessment practice. This includes further exploration of the influence of child and family factors on SLP assessment practice. Studies are also needed to determine the types of assessments being conducted by other personnel, particularly teachers and teacher-aides, for the purpose of developing a greater understanding of this practice. In addition, studies are needed to examine the reliability of having other personnel conduct assessments so that more specific recommendations can be made regarding evidence for this practice. As it is likely that SLPs experience different challenges with regards to use of different types of assessments, further investigation is needed to identify specific challenges that may be associated with use of specific types of assessments.

Surveys conducted in countries other than Australia are also needed to explore factors that may influence assessment practice at an international level. Use of the taxonomy

developed in this thesis to collect such survey data is recommended, as this would facilitate direct comparisons between Australia and other English-speaking countries.

Lastly, research is needed to explicitly investigate strategies that may assist SLPs to implement evidence-based practice recommendations for language assessment. This includes qualitative studies, such as interviews or focus groups, to further explore SLPs perceptions of factors that act as barriers and facilitators to implementation of practice recommendations and studies that measure the outcome of applying these identified strategies. Given that 'limited time' was one of the most frequently reported main challenges for SLPs in relation to assessment practice, future studies need to give particular focus to strategies that optimise the balance between quality assessment and costs of professional time.

## 7.7. Conclusion

Language assessment is an important component in service provision for school-aged children with language disorder. Therefore, it is important that SLPs use assessment practices that are evidence-based and effective in identifying the needs of children. In this thesis, important knowledge across three areas of research need was generated with the overall objective of facilitating implementation of evidence-based child language assessment practices. In Chapter 2, a systematic review was conducted to examine the psychometric properties of comprehensive norm-referenced spoken language measures. As the first systematic review of child language measures, this review added to knowledge in the field of child language assessment by summarising information from all relevant sources to make recommendations for practice. This review identified that limitations exist with regards to the psychometric evidence that is available for child language measures. In addition, this review found that available evidence is predominantly confined to studies published in assessment manuals, which limits the availability of this research evidence for SLPs. Of the comprehensive norm-referenced spoken language measures included in the review, the

Assessment of Literacy and Language (ALL), Clinical Evaluation of Language Fundamentals – 5<sup>th</sup> Edition (CELF-5), Clinical Evaluation of Language Fundamentals – Preschool: 2<sup>nd</sup> Edition (CELF:P-2) and the Preschool Language Scales - 5th Edition (PLS-5) were identified as having the most evidence and were recommended for use.

In Chapter 3, the need for detailed and agreed-upon terminology for describing child language assessment practices was addressed through development of a taxonomy with descriptive terms for describing different types of language assessments. As the first study to obtain consensus on terminology for describing child language assessments, development of this this taxonomy represents a significant step forward in the field of child language assessment. The taxonomy may be used to facilitate detailed descriptions of SLPs language assessment practice in survey research, quality assurance projects and case notes. The taxonomy also has the potential to promote much needed reflection and debate on the different types of language assessments used in clinical practice and thus may have applications in training and professional development activities.

SLPs' perceptions of factors that may influence consistent application of the taxonomy were explored in Chapter 4. Using semi-structured interviews with Delphi study participants, three main themes were identified in relation to factors that may influence application of the taxonomy: applying the taxonomy is arduous, contextual factors may influence application, and SLPs' experience and knowledge may influence application. Participants also identified practical strategies to support use of the taxonomy when collecting data from SLPs regarding the assessment practices they use. These findings from Chapter 4 provide useful information to assist with establishing consistent use of taxonomy terminology by SLPs.

In Chapters 5 and 6, a survey was used to collect data on the assessment practices Australian SLPs use when assessing the language abilities of school-aged children. Data from

this survey identified that current clinical assessment practice for school-aged children is not well-aligned with evidence-based practice. To improve assessment practice in relation to evidence-based practice, a number of recommendations are identified. Firstly, SLPs need to routinely use a range of assessments to collect information on a child's language skills in different tasks and contexts. In particular, actions are needed to increase SLPs use of assessments that are described as contextualised, activity-focussed, dynamic and targeted at school or home/community contexts. Secondly, focus needs to be placed on the importance of selecting language measures that are well-matched to the domains being assessed and the purposes for which assessment data are to be used. Thirdly, SLPs need to stop using assessments with normative data from monolingual, English speaking children with children whose background is different to this population. Fourthly, SLPs need to take actions to increase the regularity with which information is collected form parents and teachers during the assessment process. Fifthly, SLPs need to ensure that children and families are offered opportunities to receive assessment services via ICTs. Finally, SLPs should be more cautious about having teachers and teacher aides administer language assessments due to lack of evidence to support this practice.

It was identified from the survey that 'service agency', 'Australian state' and 'years of experience' may predict the regularity with which some types of assessments are used by SLPs. Frequently reported challenges related to language assessment for school-aged children were 'limited time', 'lack of assessment materials', 'limited access to training', and 'lack of confidence in assessing children from CALD backgrounds'. The most frequently reported main source of information on child language assessment was 'informal discussion with colleagues'. Collectively, these findings provide greater understanding of the context surrounding child language assessment and the factors that may influence SLPs assessment practice.

In conclusion, knowledge gained from this thesis assisted in identifying actions that need to be undertaken to improve implementation of evidence-based practice recommendations by SLPs when assessing the language abilities of school-aged children. Improving assessment practice will improve the accuracy of clinical decisions that are made from assessment data and, thus, assist in enhancing outcomes for children with language disorder.

**References for Chapter 7**

Allied Health Professions Australia. (2017). Increasing Access to Allied Health Telehealth Services. Retrieved from https://ahpa.com.au/wp-content/uploads/2017/09/170531-Telehealth-Position-Statement.pdf

American Psychological Association. (2000). *Report of the task force on test user qualifications*. Retrieved from

https://www.apa.org/science/programs/testing/qualifications.pdf

American Speech and Hearing Association. (1997-2018). *Speech language pathology assistants*. Retrieved from https://www.asha.org/Practice-Portal/Professional-Issues/Speech-Language-Pathology-Assistants/

Arias, G., & Friberg, J. (2015). Bilingual language assessment: Contemporary versus recommended practice in American schools. *Language, Speech, and Hearing Services in Schools, 48*(1), 1-15. doi:0.1044/2016_LSHSS-15-0090

Beck, A. R. (1995). Language assessment methods for three age groups of children. *Journal of Communication Development, 17*(2), 51-56. doi:10.1177/152574019501700206

Belkhodja, O., Amara, N., Landry, R., & Ouimet, M. (2007). The extent and organizational determinants of research utilization in Canadian health services organizations. *Science Communication, 28*(3), 377-417. doi:10.1177/1075547006298486

Betz, S. K., Eickhoff, J. K., & Sullivan, S. F. (2013). Factors influencing the selection of standardized tests for the diagnosis of Specific Language Impairment. *Language, Speech and Hearing Services in Schools, 44*, 133-143. doi:10.1044/0161-1461(2012/12-0093)

Bishop, D. V. M., & McDonald, D. (2009). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of*

*Language & Communication Disorders, 44*(5), 600-615.

doi:10.1080/13682820802259662

Blais, J. G., & Grondin, J. (2011). The influence of labels associated with anchor points of

likert-type response scales in survey questionnaires. *Journal of Applied Measurement,*

*12*, 370–386.

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., & Alberti, C. (2011). Using and

reporting the Delphi method for selecting healthcare quality indicators: A systematic

review. *PLoS ONE, 6*(6), e20476. doi:10.1371/journal.pone.0020476

Burke, L. A., & Hutchins, H. M. (2008). A study of best practices in training transfer and

proposed model of transfer. *Human resource development quarterly, 19*(2), 107-128.

doi:10.1002/hrdq.1230

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*, 190-200. doi:10.1044/0161-1461(2007/020)

Caesar, L. G., & Kohler, P. D. (2009). Tools clinicians use: A survey of language assessment

procedures used by school-based speech-pathologists. *Communication Disorders*

*Quarterly, 30*(4), 226-236. doi:10.1177/1525740108326334

Cameron, S., & Turtle-Song, I. (2002). Learning to write case notes using the SOAP format.

*Journal of Counseling & Development, 80*(3), 286-292. doi:10.1002/j.1556-

6678.2002.tb00193.x

Carroll, C. (2010). "It's not everyday that parents get a chance to talk like this": Exploring

parents' perceptions and expectations of speech-language pathology services for

children with intellectual disability. *International Journal of Speech-Language*

*Pathology, 12*(4), 352-361. doi:10.3109/17549500903312107

Cheung, G., Trembath, D., Arciuli, J., & Togher, L. (2013). The impact of workplace factors on evidence-based speech-language pathology practice for children with autism spectrum disorders. *International Journal of Speech-Language Pathology, 15*(4), 396-406. doi:10.3109/17549507.2012.714797

Cowie, M., Wanger, K. M., Cartwright, A., Bailey, H., Millar, J. A., & Price, M. (2001). A review of Clinical Terms Version 3 (Read Codes) for speech and language record keeping. *International Journal of Language & Communication Disorders, 36*(1), 117-126. doi:10.1080/13682820116848

Crais, E. R., Roy, V. P., & Free, K. (2006). Parents' and professionals' perceptions of the implementation of family-centered practices in child assessments. *American Journal of Speech-Language Pathology, 15*(4), 365-377. doi:10.1044/1058-0360(2006/034)

Daub, O., Skarakis-Doyle, E., Bagatto, M. P., Johnson, A. M., & Cardy, J. O. (2019). A comment on test validation: The importance of the clinical perspective. *American Journal of Speech-Language Pathology, 28*(1), 204-210. doi:10.1044/2018_AJSLP-18-0048

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical education, 40*(4), 314-321. doi:10.1111/j.1365-2929.2006.02418.x

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. *Child and Adolescent Mental Health, 20*(2), 116-125. doi:10.1111/camh.12072

Dunkley, C., Pattie, L., Wilson, L., & McAllister, L. (2010). A comparison of rural speech-language pathologists' and residents' access to and attitudes towards the use of technology for speech-language pathology service delivery. *International Journal of Speech-Language Pathology, 12*(4), 333-343. doi:10.3109/17549500903456607

Dunst, C. J., Leet, H. E., & Trivette, C. M. (1988). Family resources, personal well-being, and early intervention. *The Journal of Special Education, 22*(1), 108-116. doi:10.1177/002246698802200112

Finch, E., Cornwell, P., Ward, E. C., & McPhail, S. M. (2013). Factors influencing research engagement: Research interest, confidence and experience in an Australian speech-language pathology workforce. *BMC health services research, 13*(1), 144-155. doi:10.1186/1472-6963-13-144

Flottorp, S. A., Oxman, A. D., Krause, J., Musila, N. R., Wensing, M., Godycki-Cwirko, M., . . . Eccles, M. P. (2013). A checklist for identifying determinants of practice: A systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implementation Science, 8*(1), 1-11. doi:10.1186/1748-5908-8-35

Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact of diagnostic decisions? *Child Language Teaching and Therapy, 26*(1), 77-92. doi:10.1177/0265659009349972

Fulcher-Rood, K., Castilla-Earls, A. P., & Higginbotham, J. (2018). School-based speech-language pathologists' perspectives on diagnostic decision making. *American Journal of Speech-Language Pathology, 27*(2), 796-812. doi:10.1044/2018_AJSLP-16-0121

Gallagher, A. L., Murphy, C. A., Conway, P. F., & Perry, A. (2019). Engaging multiple stakeholders to improve speech and language therapy services in schools: An appreciative inquiry-based study. *BMC health services research, Punlished Early Online.* doi:10.1186/s12913-019-4051-z

Gillon, G., Hyter, Y., Fernandes, F. D., Ferman, S., Hus, Y., Petinou, K., . . . Westerveld, M. (2017). International survey of speech-language pathologists' practices in working

with children with autism spectrum disorder. *Folia Phoniatrica et Logopaedica, 69*(1-2), 8-19. doi:10.1159/000479063

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117-135. doi:10.1016/j.jsp.2006.05.005

Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J., Caswell, W., & Robinson, N. (2006). Lost in knowledge translation: Time for a map? *The Journal of Continuing Education in the Health Professions, 26*, 13-24. doi:10.1002/chp.47

Green, L. W., Ottoson, J. M., Garcia, C., & Hiatt, R. A. (2009). Diffusion theory and knowledge dissemination, utilization, and integration in public health. *Annual Review of Public Health, 30*, 151-174. doi:10.1146/annurev.publhealth.031308.100049

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

Hakkennes, S., & Dodd, K. (2008). Guideline implementation in allied health professions: A systematic review of the literature. *Quality and Safety in Healthcare, 17*, 296-300. doi:10.1136/qshc.2007.023804

Harding, K. E. (2014). Not enough time or a low priority? Barriers to evidence-based practice for allied health clinicians. *Journal of Continuing Education in the Health Professions, 34*(4), 224-231. doi:10.1002/chp.21255

Heilmann, J. J., Miller, J. F., & Nockerts, A. (2010). Using language sample databases. *Language, Speech, and Hearing Services in Schools, 41*(1), 84-95. doi:10.1044/0161-1461(2009/08-0075)

Hoffman, L. M., Ireland, M., Hall-Mills, S., & Flynn, P. (2013). Evidence-based speech-language pathology practices in schools: Findings from a national survey. *Language,*

*Speech, and Hearing Services in Schools, 44*, 266-280. doi:10.1044/0161-1461(2013/12-0041)

Huang, R. J., Hopkins, J., & Nippold, M. A. (1997). Satisfaction with standardized language testing: A survey of speech-language pathologists. *Language, Speech, and Hearing Services in Schools, 28*(1), 12-29. doi:10.1044/0161-1461.2801.12

Hux, K., Morris-Friehe, M., & Sanger, D. D. (1993). Language Sampling Practices: A Survey of Nine States. *Language, Speech, and Hearing Services in Schools, 24*, 84-91. doi:10.1044/0161-1461.2402.84

Kadyamusuma, M. R. (2016). Limitations to Practicing as a Bilingual Speech-Language-Pathologist in the United States. *Journal of Speech Pathology & Therapy, 1*(2). doi:10.4172/jspt.1000e103

Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *BMJ, 330*(7494), 1-8. doi:0.1136/bmj.38398.500764.8F

Kemp, K., & Klee, T. (1997). Clinical language sampling practices: Results of a survey of speech-language pathologists in the United States. *Child Language Teaching and Therapy, 13*(2), 161-176. doi:10.1177/026565909701300204

Ladyshewsky, R. K. (2010). Building competency in the novice allied health professional through peer coaching. *Journal of Allied Health, 39*(2), 77-82.

Laver, K., Brown, T., Cordier, R., & Lannin, N. (2018). Increasing research utilisation through accessible research findings for occupational therapists. *Australian Occupational Therapy Journal, 65*(4), 247-248. doi:10.1111/1440-1630.12517

LEADERS. (2014). Test Review: Clinical Evaluation of Language Fundamentals – Fifth Edition (CELF-5). Retrieved from

https://edblogs.columbia.edu/leadersproject/files/2014/02/CELF5-Test-Review-LEADERS-1-29ay7q2.pdf

Lombardino, L. J., Leiberman, R., & Brown, J. C. (2005). *Assessment of Literacy and Language*. San Antonio, TX: Pearson Psychcorp.

Marshall, J., Goldbart, J., Pickstone, C., & Roulstone, S. (2015). Application of systematic reviews in speech-and-language therapy. *International Journal of Language & Communication Disorders, 46*(3), 261–272. doi:10.3109/13682822.2010.497530

Mashima, P. A., & Doarn, C. R. (2009). Overview of telehealth activities in speech-language pathology. *Telemedicine and e-Health, 14*(10), 1101-1117. doi:10.1089/tmj.2008.0080

McCabe, P. J. (2018). Elizabeth Usher Memorial Lecture: How do we change our profession? Using the lens of behavioural economics to improve evidence-based practice in speech-language pathology. *International Journal of Speech-Language Pathology, 20*(3), 300-309. doi:10.1080/17549507.2018.1460526

Metcalfe, C., Lewin, R., Wisher, S., Perry, S., Bannigan, K., & Moffett, J. K. (2001). Barriers to implementing the evidence base in four NHS therapies: Dietitians, occupational therapists, physiotherapists, speech and language therapists. *Physiotherapy, 87*(8), 433-441. doi:10.1016/S0031-9406(05)65462-4

Michie, S., Johnston, M., Abraham, C., Lawton, R., Parker, D., & Walker, A. (2005). Making psychological theory useful for implementing evidence based practice: A consensus approach. *Quality and Safety in Healthcare, 14*(1), 26-23. doi:10.1136/qshc.2004.011155

Mokkink, L. B., Prinsen, C. A. C., Bouter, L. M., De Vet, H. C. W., & Terwee, C. B. (2015). The COnsensus-based Standards for the selection of health Measurement INstruments

(COSMIN) and how to select an outcome measurement instrument. *Brazilian Journal of Physical Therapy, 20*(2), 105-113. doi:10.1590/bjpt-rbf.2014.0143

Nail-Chiwetalu, B. J., & Ratner, N. B. (2006). Information literacy for speech-language pathologists: A key to evidence-based practice. *Language, Speech, and Hearing Services in Schools, 37*(3), 157-167. doi:10.1044/0161-1461(2006/018)

O'Callaghan, A. M., McAllister, L., & Wilson, L. (2005). Barriers to accessing rural paediatric speech pathology services: Health care consumers' perspectives. *Australian Journal of Rural Health, 13*(3), 162-171. doi:10.1111/j.1440-1854.2005.00686.x

O'Connor, S., & Pettigrew, C. M. (2009). The barriers perceived to prevent the successful implementation of evidence-based practice by speech and language therapists. *International Journal of Language & Communication Disorders, 44*(6), 1018-1035. doi:10.3109/13682820802585967

Olswang, L. B., & Prelock, P. A. (2015). Bridging the gap between research and practice: Implementation science. *Journal of Speech, Language, and Hearing Research, 58*, 1818-1826. doi:10.1044/2015_JSLHR-L-14-0305

Owen, R., Hayett, L., & Roulstone, S. (2004). Children's views of speech and language therapy in school: Consulting children with communication difficulties. *Child Language Teaching and Therapy, 20*(1), 55-73. doi:10.1191/0265659004ct263oa

Pavelko, S. L., Owens, R. E., Ireland, M., & Hahs-Vaughn, D. L. (2016). Use of language sample analysis by school-based SLPs: Results of a nationwide survey. *Language, Speech, and Hearing Services in Schools, 47*(3), 246-258. doi:10.1044/2016_LSHSS-15-0044

Pearce, W. M., & Williams, C. (2013). The cultural appropriateness and diagnostic usefulness of standardized language assessments for Indigenous Australian children.

*International Journal of Speech-Language Pathology, 15*(4), 429-440. doi:10.3109/17549507.2012.762043

Perkins, M. B., Jensen, P. S., Jaccard, J., Gollwitzer, P., Oettingen, G., Pappadopulos, E., & Hoagwood, K. E. (2007). Applying theory-driven approaches to understanding and modifying clinicians' behavior: what do we know? *Psychiatric Services, 58*(3), 342-348. doi:10.1176/ps.2007.58.3.342

Polit, D. F. (2015). Assessing measurement in health: Beyond reliability and validity. *International Journal of Nursing Studies, 52*, 1746-1753. doi:10.1016/j.ijnurstu.2015.07.002

Rappolt, S., & Tassone, M. (2002). How rehabilitation therapists gather, evaluate, and implement new knowledge. *Journal of Continuing Education in the Health Professions, 22*(3), 170-180. doi:10.1002/chp.1340220306

Ratner, N. B. (2006). Evidence-based practice: An examination of its ramifications for the practice of speech-language pathology. *Language, Speech, and Hearing Services in Schools, 37*, 257-267. doi:10.1044/0161-1461(2006/029)

Renfrew, C. (2010). *Renfrew Action Picture Test*. Camberwell, Australia: ACER.

Roulstone, S., Marshall, J., Powell, G. G., Goldbart, J., Wren, Y. E., Coad, J., . . . Coad, R. A. (2015). Evidence-based intervention for preschool children with primary speech and language impairments: Child Talk – an exploratory mixed-methods study. *Programme Grants for Applied Research, 3*(5), 1-408. doi:10.3310/pgfar03050

Ruggero, L., McCabe, P., Ballard, K. J., & Munro, N. (2012). Paediatric speech-language pathology service delivery: An exploratory survey of Australian parents. *International Journal of Speech-Language Pathology, 14*(4), 338-350. doi:10.3109/17549507.2012.650213

Samuelson, W., & Zeckhauser, R. (1998). Status quo bias in decision making. *Journal of risk and uncertainty, 1*(1), 7-59. doi:10.1007/BF00055564

Schmitt, M. B., Logan, J. A. R., Tambyraja, S. R., Farquharson, K., & Justice, L. M. (2017). Establishing Language Benchmarks for Children With Typically Developing Language and Children With Language Impairment. *Journal of Speech, Language, and Hearing Research, 60*(2), 364. doi:10.1044/2016_JSLHR-L-15-0273

Shrubsole, K., Worrall, L., Power, E., & O'Connor, D. A. (2018). Barriers and facilitators to meeting aphasia guideline recommendations: What factors influence speech pathologists' practice? *Disability and Rehabilitation*, Advance online publication. doi:10.1080/09638288.2018.1432706

Spaulding, T. J. (2012). Comparison of severity ratings on norm-referenced tests for children with specific language impairment. *Journal of Communication Disorders, 45*, 59-68.

Speech Pathology Australia. (2014). *Submission to the inquiry into the prevalence of different types of speech, language and communication disorders and speech pathology services in Australia.* Retrieved from Australia: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/Speech_Pathology

Speech Pathology Australia. (2016). *Position paper: Allied health assistants - speech pathology*. Retrieved from http://www.speechpathologyaustralia.org.au/SPAweb/Resources_for_Speech_Pathologists/Professional_Resources/HTML/Allied_Health_Assistants.aspx?WebsiteKey=fc2020cb-520d-405b-af30-fc7f70f848db

Sumsion, T. (1998). The Delphi technique: An adaptive research tool. *British Journal of Occupational Therapy, 61*(4), 153-156. doi:10.1177/030802269806100403

Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., . . . Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research, 52*(3), 643-652. doi:10.1044/1092-4388(2009/08-0136)

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, 21*(1), 10-21. doi:10.1080/2050571X.2017.1309788

Thomas-Stonell, N., Washington, K., Oddson, B., Robertson, B., & Rosenbaum, P. (2013). Measuring communicative participation using the FOCUS©: Focus on the Outcomes of Communication Under Six. *Child: care, health and development, 39*(4), 474-480. doi:10.1111/cch.12049

Thomas, E., & Magilvy, J. K. (2011). Qualitative rigor or research validity in qualitative research. *Journal for Specialists in Pediatric Nursing, 16*, 151-155. doi:10.1111/j.1744-6155.2011.00283.x

Trembath, D., Westerveld, M., & Shellshear, L. (2016). Assessing spoken language outcomes in children with ASD: A systematic review. *Current developmental Disorders Reports, 3*(33), 33-45. doi:10.1007/s40474-016-0068-8

Vallino-Napoli, L. D., & Reilly, S. (2004). Evidence-based health care: A survey of speech pathology practice. *Advances in Speech Language Pathology, 6*(2), 107-112. doi:10.1080/14417040410001708530

Van Achterberg, T., Holleman, G., Van de Ven, M., Grypdonck, M. H., Eliëns, A., & Van Vliet, M. (2006). Promoting evidence-based practice: The roles and activities of

professional nurses' associations. *Journal of Advanced Nursing, 53*(5), 605-612. doi:10.1111/j.1365-2648.2006.03751.x

Volden, J., Dodd, E., Engel, K., Smith, I. M., Szatmari, P., Fombonne, E., & Vaillancourt, T. (2017). Beyond sentences: Using the Expression, Reception, and Recall of Narratives Instrument to assess communication in school-aged children with Autism Spectrum Disorder. *Journal of Speech, Language, and Hearing Research, 60*(8), 2228-2240. doi:10.1044/2017_JSLHR-L-16-0168

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010a). Assessment of children's literacy via an internet-based telehealth system. *Telemedicine and e-Health, 16*(5), 564-575. doi:10.1089/tmj.2009.0161

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010b). Internet-based telehealth assessment of language using the CELF–4. *Language Speech and Hearing Services in Schools, 41*, 445-458. doi:10.1044/0161-1461(2009/08-0131)

Walsh, R., & IGOTF-CSD. (2006). A history of terminology: International group on terminology frameworks – communication science and disorders. . Retrieved from http://www.dhrs.uct.ac.za/sites/default/files/image_tool/images/147/History_of_CSD.pdf

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., . . . Saul, J. (2008). Bridging the gap between prevention research and practice: The interactive systems framework for dissemination and implementation. *American Journal of Community Psychology, 4*(3-4), 171-181. doi:10.1007/s10464-008-9174-z

Westerveld, M. F., & Claessen, M. (2014). Clinician survey of language sampling practices in Australia. *International Journal of Speech-Language Pathology, 16*(3), 242-249. doi:10.3109/17549507.2013.871336

Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals: Preschool - 2nd Edition* (2 ed.). Bloomington, MN: Pearson.

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals - 5th Edition*. Bloomington, MN: Pearson.

Wilson, K. S., Blackmon, R. C., Hall, R. E., & Elcholtz, G. E. (1991). Methods of language assessment: A survey of California public school clinicians. *Language, Speech and Hearing Services in Schools, 22*, 236-241. doi:10.1044/0161-1461.2204.236

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales-5th Edition* (5 ed.). Bloomington, MN: Pearson.

**Thesis Appendix A**

**Author Contribution Statement for Publication 1.**

As co-authors of the manuscript entitled *"Psychometric properties of language assessments for children aged 4-12 years: A systematic review",* we confirm that Deborah Denman has been the principal researcher and has made the following contributions*: Conceptualisation, methodology, investigation, formal analysis, writing – original draft, and project administration.

My contribution was consistent with the role of co-author and involved the following contributions*: Assistance with validation of data.

Yu-Wei Chen                    Date: 28/04/19

Our contribution was consistent with the role of supervisors and co-authors and involved the following contributions*: Assistance with conceptualisation, assistance with methodology, assistance with formal analysis, writing – review of editing, and supervision.

Reinie Cordier               Date: 26/04/19

Renée Speyer                 Date: 26/04/19

Natalie Munro                Date: 26/04/19

Wendy Pearce                 Date: 30/04/19

*Classified as per Contributor Roles Taxonomy (CRediT)

**Thesis Appendix B**

**Author Contribution Statement for Publication 2.**

As co-authors of the manuscript entitled *"Describing language assessments for school-aged children: A Delphi study"*, we confirm that Deborah Denman has been the principal researcher and has made the following contributions*: Conceptualisation, methodology, investigation, formal analysis, writing – original draft, and project administration.

Our contribution was consistent with the role of supervisors and co-authors and involved the following contributions*: Assistance with conceptualisation, assistance with methodology, assistance with formal analysis, writing – review of editing, and supervision.

Reinie Cordier          Date: 26/04/2019

Renée Speyer           Date: 26/04/2019

Natalie Munro          Date: 26/04/2019

Jae-Hyun Kim          Date: 26/04/2019

*Classified as per Contributor Roles Taxonomy (CRediT)

**Thesis Appendix C**

**Author Contribution Statement for Publication 3.**


As co-authors of the manuscript entitled *"Factors influencing speech-language therapists' application of terminology for describing paediatric language assessments",* we confirm that Deborah Denman has been the principal researcher and has made the following contributions\*: Conceptualisation, methodology, investigation, formal analysis, writing – original draft, and project administration.


Our contribution was consistent with the role of supervisors and co-authors involved the following contributions\*: Assistance with conceptualisation, assistance with methodology, assistance with formal analysis, writing – review of editing, and supervision.


Reinie Cordier            Date: 26/04/2019

Renée Speyer             Date: 26/04/2019

Natalie Munro            Date: 26/04/2019

Jae-Hyun Kim            Date: 26/04/2019

Nathan Wilson           Date: 28/04/2019


\*Classified as per Contributor Roles Taxonomy (CRediT)

**Thesis Appendix D**

**Author Contribution Statement for Publication 4.**

As co-authors of the manuscript entitled *"Language assessment for primary school children (Part I): What factors influence assessment use",* we confirm that Deborah Denman has been the principal researcher and has made the following contributions*: Conceptualisation, methodology, investigation, formal analysis, writing – original draft, and project administration.

Our contribution was consistent with the role of supervisors and co-authors and involved the following contributions*: Assistance with conceptualisation, assistance with methodology, assistance with formal analysis, writing – review of editing, and supervision.

Reinie Cordier           Date: 26/04/2019

Renée Speyer            Date: 26/04/2019

Natalie Munro           Date: 26/04/2019

Jae-Hyun Kim           Date: 26/04/2019

*Classified as per Contributor Roles Taxonomy (CRediT)

**Thesis Appendix E**

**Author Contribution Statement for Publication 5.**

As co-authors of the manuscript entitled *"Language assessment for primary school children (Part II): What reasons drive SLP choice?"*, we confirm that Deborah Denman has been the principal researcher and has made the following contributions*: Conceptualisation, methodology, investigation, formal analysis, writing – original draft, and project administration.

Our contribution was consistent with the role of supervisors and co-authors and involved the following contributions: Assistance with conceptualisation, assistance with methodology, assistance with formal analysis, writing – review of editing, and supervision.

Reinie Cordier            Date: 26/04/2019

Renée Speyer             Date: 26/04/2019

Natalie Munro            Date: 26/04/2019

Jae-Hyun Kim            Date: 26/04/2019

*Classified as per Contributor Roles Taxonomy (CRediT)

**Thesis Appendix F**

**Copyright Information for Publication 1.**

A copy of the copyright conditions for publication 1 is included here. An online version of this statement may be found at: https://www.frontiersin.org/legal/copyright-statement (Retrieved 07/06/2019).

**Thesis Appendix G**

**Copyright Information for Publication 2.**

A copy of the copyright conditions for publication 2 is included here. An online version of this statement can be found at: https://authorservices.taylorandfrancis.com/copyright-and-you/ (Retrieved 07/06/2019)

**Taylor and Francis Copyright Assignment**

In our standard author contract, you transfer – or "assign" – copyright to us as the owner and publisher of the journal (or, in the case of a society-owned journal, to that learned society).

Assigning the copyright enables us to:

- Effectively manage, publish and make your work available to the academic community and beyond.
- Act as stewards of your work as it appears in the scholarly record.
- Handle reuse requests on your behalf.
- Take action when appropriate where your article has been infringed or plagiarized.
- Increase visibility of your work through third parties.

After assigning copyright, you will still retain the right to:

- Be credited as the author of the article.
- Make printed copies of your article to use for a lecture or class that you are leading on a non-commercial basis.
- Share your article using your free eprints with friends, colleagues and influential people you would like to read your work.
- Include your article Author's Original Manuscript (AOM) or Accepted Manuscript(AM), depending on the embargo period in your thesis or dissertation. The Version of Record cannot be used. For more information about manuscript versions and how you can use them, please see our guide to sharing your work.
- Present your article at a meeting or conference and distribute printed copies of the article on a non-commercial basis.

Post the AOM/AM on a departmental, personal website or institutional repositories depending on embargo period. To find the embargo period for any Taylor & Francis journal, please use the Open Access Options Finder