Folia Phoniatrica et Logopaedica

Research Article

Folia Phoniatr Logop 2023;75:334–349
DOI: 10.1159/000530718

Received: June 23, 2022
Accepted: April 5, 2023
Published online: April 24, 2023

Standardized Measures Used Regularly by Speech-Language Pathologists when Assessing the Language Abilities of School-Aged Children: A Survey

Deborah Denman[a, b] Reinie Cordier[a, c, g] Natalie Munro[d] Jae-Hyun Kim[b] Renée Speyer[a, e, f]

[a]Curtin School of Allied Health, Faculty of Health Sciences, Curtin University, Perth, WA, Australia; [b]Department of Linguistics, Macquarie University, Sydney, NSW, Australia; [c]Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle-upon-Tyne, UK; [d]Faculty of Medicine and Health, Sydney School of Health Sciences, The University of Sydney, Camperdown, NSW, Australia; [e]Department of Special Needs Education, University of Oslo, Oslo, Norway; [f]Department of Otorhinolaryngology and Head and Neck Surgery, Leiden University Medical Centre, Leiden, The Netherlands; [g]Department of Health and Rehabilitation Sciences, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

Keywords
Language disorder · Assessment · Survey

Abstract
Introduction: This study examined speech-language pathologist (SLP)'s use of standardized language measures when assessing school-aged children. ***Method:*** A total of 335 SLPs provided information in a web-based survey regarding the standardized language measures they use for school-aged children. SLPs were asked to identify the domains targeted, purposes of use, and reasons for which regularly used standardized measures were chosen for use. ***Results:*** Findings indicated that SLPs collectively use many standardized measures, although only a small number are used regularly. SLPs reported using standardized measures to assess domains that measures are not ideally designed for and for purposes that the measures are not ideally suited to assessing. SLPs reported selecting diagnostic measures based on psychometric properties but not for screening measures. Reasons for choice varied depending on the particular measure. ***Conclusion:*** Overall, findings indicated that SLPs need to place greater focus on evidence-based practice recommendations when selecting standardized measures for use with school-aged children. Implications for clinical practice and future directions are discussed.

Introduction

Children with language disorders have persistent difficulties comprehending and producing spoken and written language compared to their peers [1]. This term includes children with developmental language disorders and children who have language disorders associated with an existing condition, such as intellectual disability. Approximately 10% of school-aged children have language disorders,

karger@karger.com
www.karger.com/fpl

Correspondence to:
Deborah Denman, Deborah.denman@mq.edu.au

placing them at high risk for academic difficulties, social-emotional problems, and behavioral difficulties [2]. To appropriately plan service needs for children with language disorders, speech-language pathologists (SLPs) must first assess a child's language abilities.

Standardized measures of language ability have a structured process for evaluating language abilities, including specific guidelines for administration and scoring [3]. Standardized measures may also have norm-referenced scores which allow for a child's abilities to be compared to a sample of age-matched peers [4]. In contrast, non-standardized procedures do not have set guidelines for administration and scoring, for example, unstructured clinical observations, judgments from a sample of the child's language, and interviews with parents and caregivers [3]. Both standardized measures and non-standardized procedures provide important data on language abilities. Nonetheless, findings from previous studies of SLP practice indicate that SLPs do rely predominantly on the use of standardized measures when making diagnostic decisions [5, 6] and when planning service provision [7]. In addition, school settings often require standardized measures to be used when determining a child's eligibility for services or funding [8, 9].

Given the frequency with which standardized language measures are used and the importance of assessment results appropriated to decision-making following the assessment process, it is critical to identify which standard measures are used most regularly by SLPs. Knowing the measures that are used most regularly by SLPs is important for developing future actions to support SLPs' professional practice in relation to their actual clinical practice. In addition, by first identifying the specific measures that SLPs use most regularly, it is then possible to examine whether their use of these measures is aligned with evidence-based practice recommendations [9–11]. If SLPs are not making sound decisions regarding the measures that they use regularly when assessing school-aged children, then this has significant clinical implications for the majority of school-aged children undergoing language assessment. For the purpose of this current study, regular use of a standardized measure was defined as the measure being used with half or more of the last 40 children on the SLP's caseload. This definition is consistent with the criteria used in a previously published study reporting on factors that influence SLP assessment practice [12].

Evidence-Based Practice when Selecting Standardized Language Measures

Standardized language measures for school-aged children may target a range of domains, for example, semantics, morphosyntax, social abilities and discourse, meta-abilities, and executive functioning [3]. Each domain is best assessed using measures that were specifically designed for measuring the domain of interest. For example, evidence-based practice recommendations identify that measures targeting social abilities and discourse should include tasks that examine language in situations that are reflective of naturalistic, everyday communication environments [13–15]. If a measure is used to make judgments about domains that the measure was not designed to assess, then a child's needs may not be accurately identified [10].

Language assessment may also be conducted for a range of different purposes, for example, predicting outcomes, selecting intervention goals, determining intervention dosage, screening, diagnosis, detecting change, or describing status [3]. To align with evidence-based practice, standardized measures need to be well matched to the purposes for which data will be used [9, 10, 16]. For example, measures used for screening and diagnostic purposes should have sound psychometric properties (including internal consistency, reliability, measurement error, content validity, criterion validity, structural validity, and hypothesis-testing), as well as sound diagnostic accuracy; measures used for detecting change should have evidence of being responsive to meaningful change in functional performance; and measures used for selecting intervention goals should target performance in real-life situations [13, 16–18]. Using a measure for purposes for which the measure was not ideally designed may compromise the soundness of decisions made from the assessment data [10, 14].

SLPs may consider a range of factors when selecting standardized language measures for use with school-aged children, including practical factors (e.g., availability, time, and cost), policy factors (e.g., workplace regulations), and experiential factors (e.g., familiarity with measure, peer suggestions) [11]. Nonetheless, a strong focus on scientific evidence needs to be maintained when SLP chooses standardized measures for use, as this is not compensated for by other factors. For example, a measure that lacks evidence for diagnostic accuracy is not suitable for diagnostic purposes regardless of practical, policy, or experiential factors or a child's linguistic profile [6]. Understanding SLPs' reasoning behind the selection of standardized language measures is important for effectively identifying and addressing potential barriers to the implementation of evidence-based practice recommendations.

Standardized Measures Used by SLPs

To date, most studies of SLP language assessment practice for school-aged children have focused on surveys

of SLPs in the USA and UK [9, 11, 19]. These studies have identified that the most commonly used standardized language measures for school-aged children are comprehensive language measures or single-word vocabulary measures [9, 11, 19]. From this, it has been identified that SLPs appear to favor measures with normative data from the country in which they work. For example, measures developed in the UK such as the Renfrew Action Picture Test (RAPT) [20] and editions of the Reynell Developmental Language Scales [21] are commonly used by SLPs in the UK but not in the USA [9, 22].

These previous studies identify that SLPs may not always give sufficient consideration to the psychometric quality of measures in relation to the purpose for which data are used, for example, by ensuring that measures used for diagnostic purposes have established diagnostic accuracy [6, 9]. It has also been identified that SLPs may not be choosing measures that best match the domains that need to be assessed, for example, by overusing single-word vocabulary measures when measures that target other domains, such as morphology, syntax, or discourse, may be more appropriate for diagnosis and intervention planning [11].

Despite the widespread and frequent use of standardized language measures, only two previous studies have investigated the reasons that SLPs report for selection of standardized measures [6, 11]. Findings identified that SLPs do consider a range of factors when selecting measures for use, although ease of administration/scoring and availability/familiarity with the measure were the most frequently reported reasons for selection of a standardized measures [11]. This is also consistent with data showing that SLPs appear to favor familiar measures that have longstanding histories in the field with multiple editions, such as five editions of the Clinical Evaluation of Language Fundamentals (CELF) [23–25].

These earlier studies have provided important information regarding SLPs' selection of standardized measures as a general category, i.e., when contrasted with reasons for choosing non-standardized procedures. No previous studies have specifically focused on investigating the decisions SLPs make when selecting a particular standardized measure for use based on the language domains, the purposes of assessment, and the main reasons for choosing the measure. Therefore, it is not known how SLPs are using particular standardized measures and whether SLPs' decision-making aligns with evidence-based practice recommendations. To identify the extent to which SLP's use of standardized measures aligns evidence-based practice, information is needed regarding (1) the domains that SLPs target when using a particular standardized language measure (e.g., morphology, syntax, discourse), (2) the purposes for which data from a particular language measure is used (e.g., diagnosis, screening, intervention planning), and (3) the reasons why a particular measure is chosen for use (e.g., psychometric properties, ease of administration). This information will assist in identifying future actions that may be needed to support implementation of evidence-based language assessment practices for school-aged children.

Aims

The aims of this study were to answer the following research questions.

1. What particular standardized measures do SLPs regularly use to assess the language abilities of children aged 4–12 years?
2. What are the main domains targeted, the main purposes of use, and the main reasons why regularly used standardized measures are chosen?

The aim of this study was to investigate the standardized measures that SLPs use regularly when assessing school-aged children and the domains assessed, purposes of use, and reasons for choice in relation to regularly used standardized measures. Therefore, a broad population of pediatric SLPs was chosen as a focus for this study. Information on SLP practice at a broad level is important as it provides data on practice trends and highlights areas of practice that may benefit from further investigation in future studies. This study was conducted in one country (Australia); however, findings from this survey have relevance to SLPs internationally.

Method

The Checklist for Reporting Of Survey Studies (CROSS) was used in the reporting of the methodology in this study [26].

Survey Design

This study was designed as a cross-sectional survey. The online survey was created using Qualtrics software [27], and questions were developed from literature on survey design [28, 29]. Prior to dissemination of the survey, four SLPs working as clinicians trialed the survey concurrently to confirm clarity of questions. To ensure diverse representation, the pilot SLPs were all from different service agencies, including private practice, public education (school) department, non-government disability service agency, and a university teaching clinic. Minor edits were made to the formatting in response to pilot feedback. Survey completion time was estimated at between 25 and 40 min 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.

Survey Structure and Format

The survey consisted of four sections. Section one asked questions to establish eligibility to complete the survey. Respondents that did not indicate that they provided clinical services to at least 40 children aged 4–12 years with oral or written language disorders in the previous 12 months were not able to continue with the survey. Section two collected data on participant demographics. Questions were multiple choice with "other" options available where applicable. Data from section three related to different research aims which are reported in a separate manuscript [12]. Section four is relevant to this study and was designed to elicit responses from participants regarding the standardized language measures they used for the last 40 children who accessed their services. This included questions about the names of specific standardized measures used and regularity with which each measure was used. See Figure 1 for a summary of the survey structure and format.

In section four, participants were first asked to list the names of standardized measures they used in open response boxes and then indicate the regularity with which each measure was used on a 5-point Likert scale. Data from these questions addressed the first study aim. To facilitate consistent application of the frequency rating scale by survey participants, Likert scale points were associated with numeric qualifiers as well as descriptive terms [28]. Participants were asked: "How many children were assessed using the standardized measure 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). SLPs were asked to consider the last 40 children in survey questions as it was considered that 40 was a large enough number to capture trends but not so large that participants would be unable to accurately recall the standardized measures they used.

For standardized measures that participants identified as using regularly, participants were then asked to list the main domains (up to three), main purposes (up to three) for which each measure was used, and main reasons why the measure was chosen for use (up to three). Data from these questions addressed study aim two. Investigating the domains assessed, main purposes, and reasons for selection of measures that are used regularly (i.e., with "many" or "most" children) was a priority given the number of children potentially affected if SLPs are not making evidence-based decisions. This criterion also facilitated accurate reporting by survey participants (i.e., it is likely easy for SLPs to accurately recall measures that were used with more than half of children). Asking participants to only report on measures that they used with half or more children also made the survey completion time manageable enough to ensure a high response rate and reduced the likelihood of responses being influenced by participants being unfamiliar with some measures (i.e., as SLPs were not being asked to comment on measures that they may have only used a small number of times).

Questions were multiple choice with open response boxes where appropriate. For example, participants were asked, "what domains do you primarily focus on when you use the standardized assessment?" Participants were restricted to selecting three main domains, purposes, and reasons for each assessment as the research objective was to investigate the main or primary domains and purposes for which measures were used. Allowing participants to select more than this number of options may have resulted in participants selecting domains and purposes that were not main or primary domains, purposes, or reasons.

To ensure SLPs applied the same definitions when describing assessment domains and purposes, terms and definitions from a recently developed taxonomy were used in the survey questions [3]. 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. The taxonomy terms and definitions that were used in survey section 4 are provided in online supplementary material 1 (for all online suppl. material, see https://doi.org/10.1159/000530718). 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. A copy of the survey questions is provided in online supplementary material 2.

Survey Dissemination

The survey was accessible between mid-February and mid-June 2018. All SLPs working in Australia who met eligibility criteria were permitted to participate. The link to the survey was distributed through Twitter, Facebook 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 it around their professional networks. The online survey software was programmed to only allow one survey to be completed from each IP address.

Survey Participants

In total, 847 survey responses were received, with 727 being complete and valid survey responses (85.8% total completion rate). Of the SLPs who completed the survey, 525 SLPs identified themselves as working with children 4–12 years with language disorders. Using membership data from the Australian association for speech pathologists, Speech Pathology Australia, the 525 responses in this survey represented approximately 11.4% of the estimated target population size at the time of the survey.

Participant Characteristics

Of the 525 SLPs who indicated in the survey that they worked with school-aged children, 407 SLPs indicated having provided a service to 40 or more children with language disorders in the preceding year. Of these 407, 335 SLPs completed all the survey questions in section 4 of the survey (82.3% completion rate for SLPs who indicated regularly providing services to school-aged children with language disorders). Participant characteristics reported by these SLPs are outlined in Table 1. Data from participants who did not fully complete section four were not included. 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 disorders but did not complete survey section 4 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.

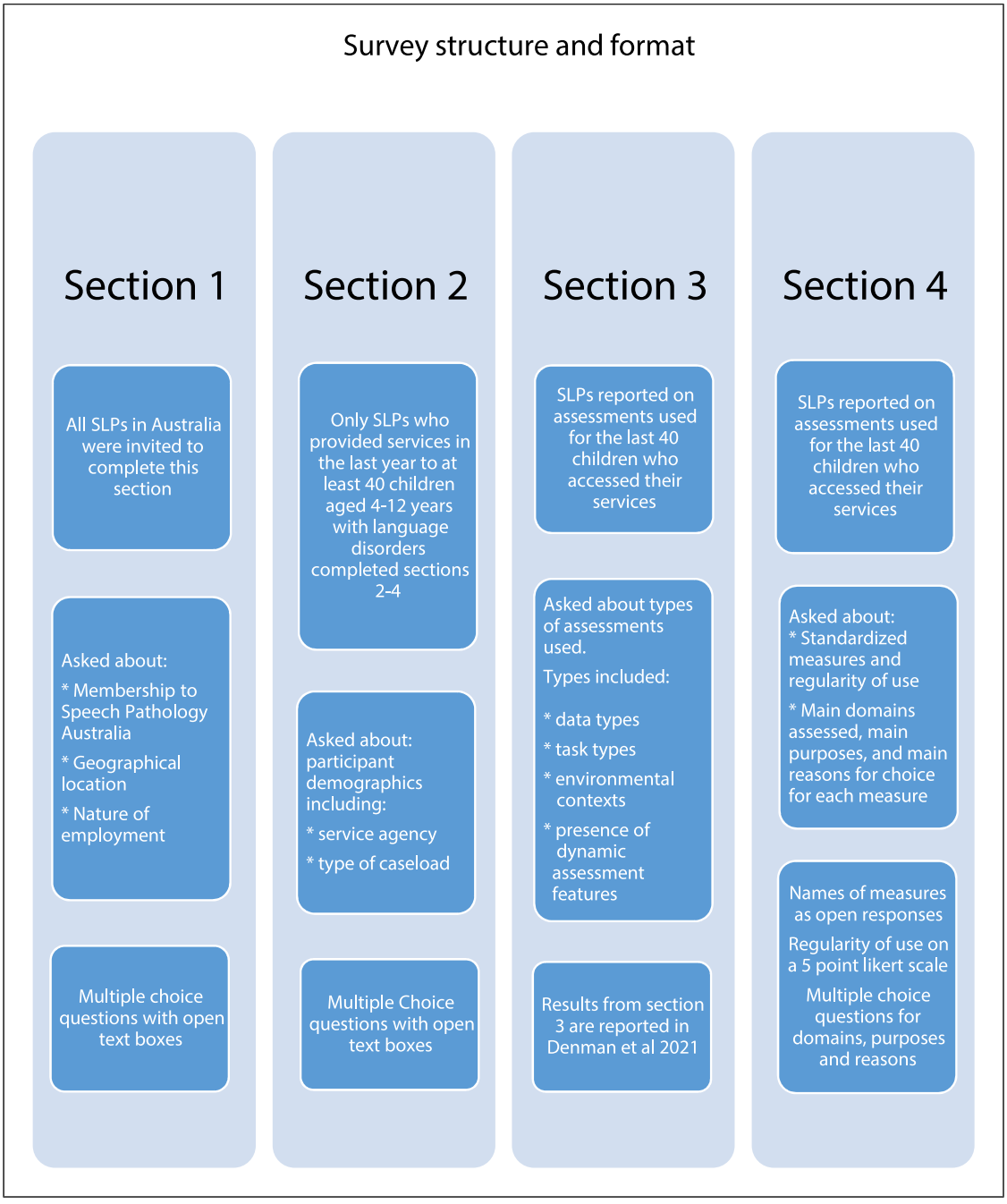

Fig. 1. Sections within the survey.

Data Analysis

Data from the survey was exported from Qualtrics into the Statistical Package for the Social Sciences (SPSS) version 20 program [30]. Descriptive statistics were used to report on the regularity with which different standardized measures were used by SLPs and the domains, purposes, and reasons for which SLPs reported using specific standardized measures. For the purposes of survey analysis, "regular use" was defined as selection of Likert scale points for "many" or "most" children (i.e., with half or more than half of the last 40 children who received services).

To analyze data from the question about standardized measures, different editions of the same measure were combined during analysis. This was because SLPs typically used only one edition of a standardized measure and did not always specify the edition used. The exceptions were editions of the CELF [25] and the Test of Narrative Language [31]. As some SLPs identified using two

Table 1. Demographics of participants who provided services to at least 40 children with language disorder in last year and completed the survey (n = 335)

Category	Subcategory	Total (%)
Gender	Female	323 (96.4)
	Male	11 (3.3)
	Other	1 (0.3)
	Total	**335 (100)**
Australian State/territory	Australian Capital Territory (ACT)	7 (2.1)
	New South Wales (NSW)	79 (23.6)
	Northern Territory (NT)	7 (2.1)
	Queensland (QLD)	101 (30.1)
	South Australia (SA)	31 (9.3)
	Tasmania (TAS)	13 (3.9)
	Victoria (VIC)	55 (16.4)
	Western Australia (WA)	42 (12.5)
	Total	**335 (100)**
Agency through which service is provided	Health agency[1]	50 (14.9)
	Education agency[2]	117 (34.9)
	Private practice[3]	104 (31.0)
	Disability specific agency[4]	43 (12.8)
	General agency[5]	15 (4.5)
	University clinic[6]	6 (1.8)
	Total	**335 (100)**
Remoteness of geographical location[7]	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 CALD backgrounds	High frequency of CALD[8]	52 (15.5)
	Low frequency of CALD[9]	283 (84.5)
	Total	**335 (100)**

CALD, culturally and linguistically diverse. [1]Government or non-government health service or hospital. [2]Government or non-government education service or school. [3]Private practice, i.e., business owner or employee in private practice. [4]Government or not government agency with eligibility criteria stipulating that children must have diagnosis (or suspected diagnosis) of disability to access service. [5]Government or non-government agency – not identified as education, health, or disability. [6]University clinic (e.g., teaching clinic). [7]Determined by classifying work postcode according to 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. [8]More than half of children with language disorder on caseload were from CALD backgrounds, e.g., bilingualism or standard Australian English is not the first language. [9]Less than half of children with language disorder on caseload were from CALD backgrounds, e.g., bilingualism or Standard Australian English is not the first language.

different editions of these measures, different editions of these particular measures were counted as two different measures. No SLPs identified the TNL or CELF without supplying the edition. One participant indicated Clinical Evaluation of Language Fundamentals: Preschool (CELF-P) and this was combined with responses for CELF-P:2. Any responses listed by SLPs in response to the question about the standardized measures that they used that were not measures of language, for example, speech production measures, were removed from analysis. A total of 50 (2.2%) responses were also removed from the analysis as they did not refer to the names of specific standardized measures. This included broad descriptors such as "language sampling" or "language screener" or acronyms that were ambiguous or could not be identified in online Google searches.

Results

Standardized Measures Used Regularly by SLPs

The 335 SLPs who completed the survey collectively listed 139 identifiable standardized measures as being used

Table 2. Standardized language measures used by SLPs and regularity of use (n = 335)

Standardized language measure (i.e., defined guidelines for administration and scoring)	Author and publication year	Number of SLPs who used the assessment, n (%)	Number of SLPs who used the assessment regularly,* n (%)
CELF-4 Core and Language Index Subtests (Clinical Evaluation of Language Fundamentals – fourth Edition)	Wiig et al. [24]	245 (73.1)	125 (37.3)
CELF:P Core Language Subtests (Clinical Evaluation of Language Fundamentals Preschool – any edition)	Wiig et al. [23]	222 (66.2)	46 (13.7)
RAPT (Renfrew Action Picture Test – any edition)	Renfrew [20]	214 (63.8)	93 (27.8)
CELF-5 Core and Language Index Subtests (Clinical Evaluation of Language Fundamentals – fifth Edition)	Wiig et al. [25]	191 (57.0)	60 (17.9)
SPAT (Sutherland Phonological Awareness Assessment – any edition)[1]	Neilsen [32]	191 (57.0)	42 (12.5)
PLS (Preschool Language Scales – any edition)	Zimmerman et al. [33]	138 (41.1)	8 (2.4)
CELF-4 Working Memory Subtests (Clinical Evaluation of Language Fundamentals – fourth Edition)	Wiig et al. [24]	65 (19.4)	7 (2.1)
TNL (Test of Narrative Language)[2]	Gillam and Pearson [34]	57 (17.0)	5 (14.9)
CELF-5 Reading Comprehension and Structured Writing Subtests (Clinical Evaluation of Language Fundamentals – fifth Edition)[1,2]	Wiig et al. [25]	52 (15.2)	3 (0.1)
CELF-5 Pragmatic Profile (Clinical Evaluation of Language Fundamentals – fifth Edition)[2,3]	Wiig et al. [25]	51 (15.2)	7 (2.1)
TOPS (Test of Problem Solving) (elementary or adolescent – any edition)	Bowers et al. [35]	47 (14.1)	3 (0.1)
CELF-4 Pragmatics Profile (Clinical Evaluation of Language Fundamentals – fourth Edition)[2,3]	Wiig et al. [24]	36 (10.7)	3 (0.1)
YARC – Passage Reading (York Assessment of Reading for Comprehension – any edition)[1,2]	Snowling, Stothard [36]	35 (10.4)	7 (2.1)
RBS (Renfrew Bus Story – any edition)[2]	Renfrew [37]	33 (9.9)	8 (2.4)
CCC (Children's Communication Checklist – any edition)[3]	Bishop [38]	32 (9.5)	4 (1.2)
NARA (Neale Analysis of Reading Ability – any edition)[1,2]	Neale [39]	32 (9.5)	3 (0.1)
PPVT (Peabody Picture Vocabulary Test – any edition)	Dunn and Dunn [40]	31 (9.3)	4 (1.2)
Reynell (Reynell Developmental Scales – any edition)	Edwards et al. [41]	28 (8.45)	1 (0.3)
CTOPP (Comprehensive Test of Phonological Processing – any edition)[1]	Wagner [42]	24 (7.2)	7 (2.1)
QUIL (Queensland Inventory of Literacy)[1]	Dodd and The University of Queensland Dept. of Speech Pathology and Audiology [43]	22 (6.6)	0 (0.0)
YARC – Early Reading (York Assessment of Reading for Comprehension)[1]	Snowling, Stothard [36]	22 (6.6)	1 (0.3)
CELF-4 Screening Test (Clinical Evaluation of Language Fundamentals Screening Test – fourth Edition)	Semel et al. [44]	21 (6.3)	7 (2.1)
TNL −2 (Test of Narrative Language – second Edition)[2]	Gillam and Pearson [31]	21 (6.3)	1 (0.3)
Communication Matrix[3]	Rowland and Fried-Oken (2004) Rowland and Fried-Oken [45]	20 (6.0)	1 (0.3)
TOLD-I (Test of Language Development – Intermediate – any edition)	Newcomer and Hammill [46]	19 (5.7)	0 (0.0)
CASL (Comprehensive Assessment of Spoken Language)	Carrow-Woolfolk [47]	18 (5.4)	2 (0.6)
ERRNI (Expression, Reception, Recall of Narrative Instrument)[2]	Bishop [48]	16 (4.8)	0 (0.0)
OWLS-II (Oral and Written Language Scales - 2nd Edition) listening and/or speaking components	Carrow-Woolfolk [49]	15 (4.5)	0 (0.0)

Table 2 (continued)

Standardized language measure (i.e., defined guidelines for administration and scoring)	Author and publication year	Number of SLPs who used the assessment, *n* (%)	Number of SLPs who used the assessment regularly,* *n* (%)
TOLD-P (Test of Language Development – Primary – any edition)	Hammill and Newcomer [50]	15 (4.5)	0 (0.0)
Bureau Test of Auditory Comprehension	Health Commission of New South Wales [51]	15 (4.5)	3 (0.1)
Peter and the Cat Retell[2]	Leitao and Allan [52]	13 (3.8)	1 (0.3)
PLAI (Preschool Language Assessment Instrument)	Blank et al. [53]	13 (3.8)	1 (0.3)
TOPL-2 (Test of Pragmatic Language - Second edition)	Phelps-Terasaki and Phelps-Gunn [54]	13 (3.8)	0 (0.0)
CELF:P Pragmatics Profile (Clinical Evaluation of Language Fundamentals: Preschool – second Edition)[2,3]	(Wiig et al., 2004) [23]	11 (3.3)	1 (0.3)
OWLS-II Reading and Writing Tests (Oral and Written Language Scales - 2nd Edition)[1,2]	Carrow-Woolfolk (2011) Carrow-Woolfolk [49]	10 (3.0)	1 (0.3)
YARC (York Assessment of Reading for Comprehension) – not otherwise specified[1]	Snowling, Stothard [36]	10 (3.0)	0 (0.0)
ONAP (Oral Language Assessment Package): DECD South Australia[2]	Government of South Australia: Department of Education and Children's Services [55]	9 (2.7)	1 (0.3)
South Australian Spelling Test[1]	Westwood [56]	9 (2.7)	1 (0.3)

*Regular use was identified if a language measure or procedure was used with 20 or more of the last 40 children who received services (i.e., half or more than half of the last 40 children on SLP caseload). Language measures or assessment procedures that were listed by 2.4% (8/335) or less participants are not included in this table. [1]Language measure that is known to target written language. [2]Language measure that is known to target social abilities and discourse. [3]Language measure or procedure that uses proxy-reporting (i.e., information reported by others such as a checklist or questionnaire).

to assess the language abilities of children. On average, each SLP listed 6.9 (SD = 3.23) different standardized measures. Six participants indicated not using any standardized measures (these SLPs used only non-standardized assessment procedures), and two SLPs identified using 15 or more different standardized measures. A list of standardized measures used by SLPs is provided in Table 2.

Only five measures were identified as being used regularly (i.e., with half or more of the last 40 children on the SLP's caseload) by more than 2.4% (8/335) of SLPs who participated in the survey. These five measures were the CELF – 4th Edition (CELF-4) core or language index subtests (used regularly by 37.3% or 125/335 SLPs) [24], the CELF – 5th Edition (CELF-5) core or language index subtests (used regularly by 17.9% or 60/335 SLPs) [25], the CELF-P – any edition core language subtests [23], the RAPT – any edition (used regularly by 27.8% or 93/335 SLPs) [20], and the Sutherland Phonological Awareness Test – any edition (SPAT; used regularly by 12.5% or 42/335 SLPs) [32].

The CELF-4 and CELF-5 measures are designed for children aged 5–21 years. The core and language index subtests are administered by an SLP under test-taking conditions. The subtests include word and sentence level tasks that target language comprehension or production of semantics, morphology, and/or syntax [24, 25]. The CELF-P is designed for children aged 3–6 years. Similarly to the CELF-4 and CELF-5, the core and language index subtests are administered by an SLP under test-taking conditions and include word and sentence level tasks targeting language comprehension or production of semantics, morphology, and/or syntax [23]. The RAPT is designed for children 3–8 years of age. The measure is administered by an SLP or teacher under test-taking conditions and consists of 10 picture cards with a question accompanying each question to elicit a sentence from the child. The RAPT targets expressive semantics, syntax, and morphology [20]. The SPAT has normative data for Australian children in the first 4 years of schooling. The measure is administered by an SLP or teacher under

test-taking conditions. The SPAT consists of a series of tasks designed to assess phonological and phonemic awareness skills at single-word level. These tasks include rhyming, sound blending, sound segmenting, sound manipulation, non-word reading, and non-word spelling [32].

Given that these measures were widely used across the collective population of SLPs, there are significant potential consequences if these particular measures are not used appropriately. All other standardized measures were used regularly by less than eight SLPs, which is not a sufficient sample size from which to draw sound conclusions about SLP practice more generally. For these reasons, this manuscript focuses specifically on these five standardized measures.

Domains, Purposes, and Reasons in Relation to Regularly Used Standardized Measures

Most SLPs reported targeting the domains of semantics and morphosyntax when using the CELF core language or language index subtests. However, almost one-quarter of SLPs also selected a focus on social abilities and discourse when they used the CELF-4 core language or language index subtests. The RAPT was primarily used to assess morphosyntax and semantics. Approximately one quarter of SLPs who regularly used the RAPT selected social abilities and discourse as a primary domain assessed by the RAPT. Most SLPs used the SPAT to assess meta-abilities, although approximately one-fifth to one-sixth of SLPs also selected executive functions or morphosyntax as areas they primarily focus on when using this measure. The main domains for which regularly used standardized measures were used to assess are shown in Figure 2.

In relation to the purpose of measures, over 80% of SLPs who regularly used the CELF core language or language index subtests reported using these measures for diagnostic purposes and approximately half indicated using these measures for purposes of predicting outcome, selecting intervention, or describing status. Similarly, over 80% of SLPs who regularly used the RAPT reported using this measure for screening purposes and approximately half indicated using this measure for the purposes of detecting change. The SPAT was reported as being predominantly used for predicting outcomes and selecting interventions. The purposes for which regularly used standardized measures were used can be found in Table 3.

When the top three reasons for use of each standardized measure were combined, the most frequently identified reason for each of the three CELF core or language index subtests was "presence of Australian norms." This reason was selected by half to one-third of SLPs who regularly used the CELF-4, CELF-5, or CELF-P core or language index subtests. Over one-third of SLPs also selected "good psychometrics" and "good for selecting goals" as reasons for use of the CELF measures.

The most frequently identified reason for use of the RAPT was "quick to administer," with over 90% of SLPs who regularly used the RAPT selecting this as a reason for doing so. Almost three-quarters of SLPs who regularly use the RAPT selected "quick to score" as a reason for use of the RAPT and one third selected "good for selecting goals" or "good for selecting intervention." The most frequently identified reason for use of the SPAT was "good for selecting goals" with half of SLPs who regularly used the SPAT, selecting this as a reason for doing so. Other frequently selected reasons for use of SPAT included "quick to score," "recently developed norms," and "quick to administer" with more than one third of SLPs selecting these reasons. The reasons for which frequently used standardized measures were chosen for use are displayed in Table 4.

Discussion

Standardized Measures Used Regularly by SLPs

This study investigated the standardized measures that a broad population of SLPs use to assess the language abilities of school-aged children. Findings from this survey indicate that, although SLPs collectively use a great number of different standardized measures, only a small number are chosen for use on a regular basis (i.e., with half or more children). These measures included CELF-4, CELF-5, CELF:P-2, RAPT, and SPAT. Given the wide use of these measures, it is of particular importance that the use of these measures be examined in relation to evidence-based practice. All five of the most regularly used standardized measures by SLPs in this survey were norm-referenced measures administered under test-taking conditions to assess language at word and sentence level. As reported in a previous study, SLPs most regularly use norm-referenced measures that are de-contextualized in nature and less regularly use tasks that are more contextualized, such as language sampling [12]. In addition to Australian normative data, the versions of the CELF have normative data from the USA and UK and have been reported in previous studies as being commonly used in these countries [9, 22]. The RAPT has normative data from the UK and has been identified as frequently used in the UK [22, 57]. A different finding from this survey compared to previous surveys was the low reported use of standardized

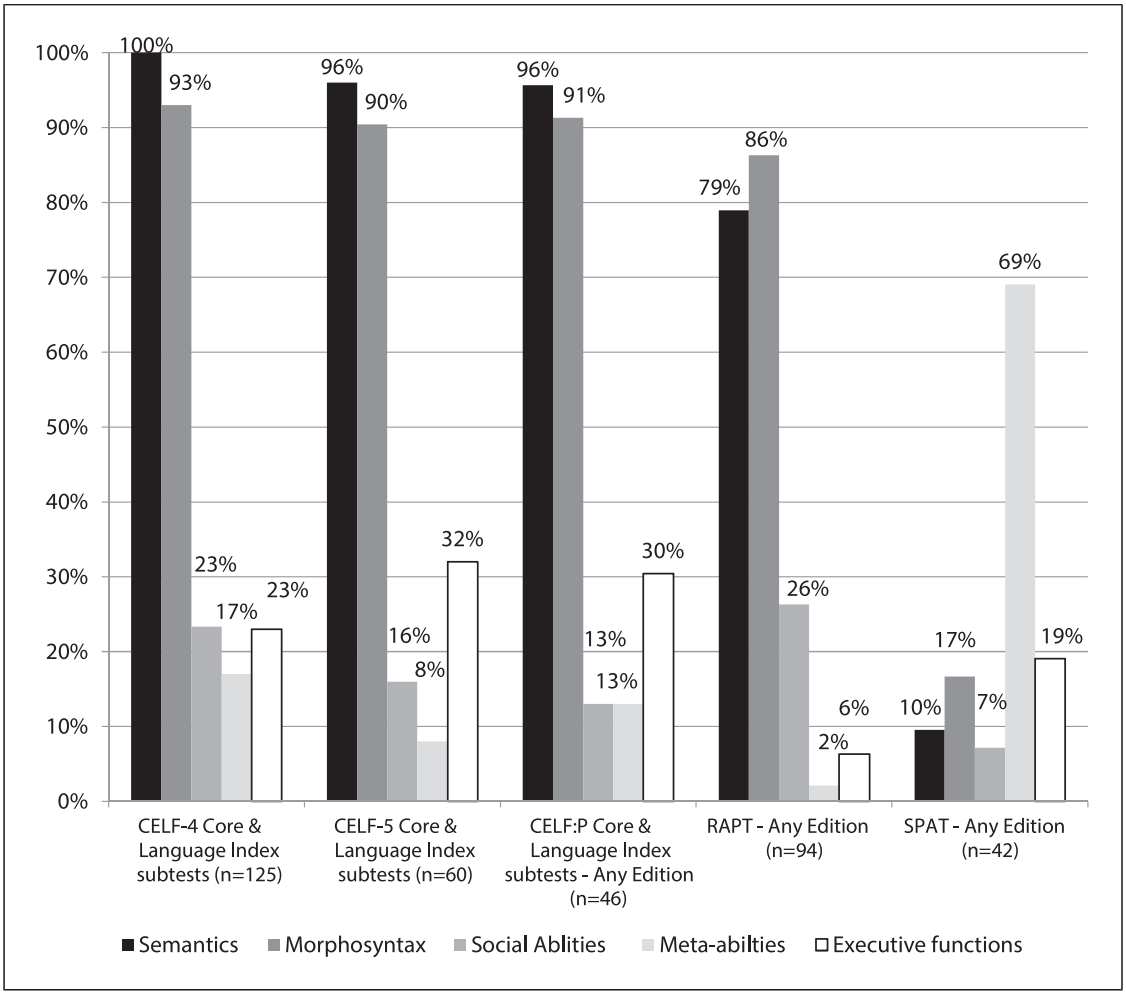

Fig. 2. Domains targeted by the five most regularly used standardized measures (n = 335). Standardized measures have set guidelines for administration and scoring. Regular use of a measure was defined if the measure was used with half or more than half of the last 40 children. The sample size (n) is different for each measure as not all SLPs used all measures regularly. SLPs could select up to three main domains they target for each measure. On average, SLPs selected 2.2 domains for each measure.

single-word vocabulary measures. This contrasts with previous surveys of SLP assessment practice in the USA which identified single-word vocabulary measures as being frequently used [9, 11, 19]. This difference may be due to a lack of single-word vocabulary measures in accordance with Australian norms. However, this finding could also reflect a more recent positive shift away from use of measures that solely target single-word vocabulary. Scores from single-word vocabulary measures are typically not considered indicative of a child's performance in naturalistic tasks or everyday communication [58, 59] and have not been shown to have better identification accuracy than measures that assess language at sentence or discourse level [11, 18].

Measures with normative data from monolingual speaking populations are not appropriate for use when making diagnostic decisions about the language abilities of children from culturally and linguistically diverse (CALD) backgrounds [60]. Previous survey research has reported that SLPs with a higher proportion of children from CALD backgrounds on their caseloads did not report less regular use of norm-referenced measures [12]. Given that the most regularly used standardized measures identified in this study were also norm-referenced, it is important that SLPs place focus on making evidence-based decisions when choosing assessments for use with children from CALD backgrounds [61].

Table 3. Main purposes for which regularly used standardized language measures were used

Purposes	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), %
Predict outcome	55.2	61.7	63.0	26.9	59.5
Select intervention	47.2	48.3	58.7	31.2	64.3
Plan dosage	4.0	1.7	4.3	1.1	2.4
Diagnosis	83.2	86.7	84.8	14.0	38.1
Screening	1.6	1.7	13.0	80.6	42.9
Detect change	41.6	28.3	28.3	47.3	38.1
Describe status	52.8	53.3	41.3	34.4	33.3

Standardized measures have set guidelines for administration and scoring. Regular use of a measure was defined as being used with half or more than half of the last 40 children. The sample size (n) is different for each measure as not all SLPs used all measures regularly. This table shows 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 each measure, i.e., a purpose was taken to be a "main purpose" if it was one of the (up to three) purposes selected by an SLP. On average, SLPs selected 2.7 purposes for each measure. See Table 1 for definitions of terms for describing purposes.

Domains, Purposes, and Reasons in Relation to Regularly Used Standardized Measures

This was the first survey to examine the domains targeted, purposes, and reasons for which particular standardized language measures are used. Findings indicate that some aspects of SLPs' use of standardized measures may not be well aligned with evidence-based practice recommendations, with resulting implications for clinical service provision. A notable finding from this survey was that one quarter of SLPs reported that social abilities and discourse are the main areas they assess when using the CELF core or language index subtests and the RAPT, despite these measures not being designed for making judgments on these domains [22]. The CELF core and language index subtests and the RAPT consist of structured table-top tasks involving interaction with an adult under test-taking conditions. These measures do not allow for language abilities to be observed in a natural social context, which is important for making accurate judgments on social abilities [14, 62]. Furthermore, both the CELF core and language index subtests and the RAPT of primarily consist of word- or sentence-level tasks which are not sufficient for assessing discourse abilities [63]. Using word- and sentence-level measures administered under test-taking conditions to make judgments about social abilities and discourse may lead to under-identification of language discourse difficulties in school-aged children [63]. To align with evidence-based practice recommendations, it is important that social abilities and discourse are assessed using text-level tasks such as language sampling [63] and/or data on performance in daily activities, for example, observations of a child's interactions in the classroom, parent/teacher interviews, or questionnaires [64].

The core or language index tests in the CELF have evidence of robust psychometric quality [65]; therefore, the finding in this survey that SLPs use these measures for diagnostic purposes is positive. It is also encouraging to note that over one-half of SLPs reported that Australian norms were a main reason for choosing of the CELF measures. This is a positive finding, although SLPs should be aware that this should not be the leading factor when selecting standardized measures, as having normative data on the population of interest does not mean measures have sound reliability and validity [65].

Besides diagnostic purposes, SLPs in this survey also reported using the CELF core or language index subtests for purposes of predicting outcome, selecting intervention, and detecting change. In addition, one third of SLPs identified "good for selecting intervention goals" as one of the main reasons the CELF core or language index subtests were chosen for use. Norm-referenced measures such as the CELF core or language index subtests are specifically designed to measure performance in relation to peers. These measures may not reflect performance in real life and may not assess each language target systematically or in enough depth to adequately determine which targets should be

Table 4. Main reasons for which regularly used standardized language measures were chosen for use

Reasons	CELF-4 core or language index (n = 125), %	CELF-5 core or language index (n = 60), %	CELF-P core or language index (n = 46), %	RAPT – any edition (n = 94), %	SPAT – any edition (n = 42), %
Australian norms	**61.6**	**56.7**	**56.5**	4.3	23.8
Good psychometrics	**40.0**	**36.7**	**37.0**	2.1	11.9
Employer requires use of assessment	**33.6**	**38.3**	19.6	2.1	9.5
Good for selecting goals	32.0	36.7	34.8	**38.3**	**50.0**
Referring agent requires use of assessment	24.0	13.3	8.7	1.1	0.0
Only available Assessment for purpose	21.6	13.3	17.4	5.3	11.9
Quick to administer	16.8	5.0	**39.1**	**91.5**	38.1
Good for selecting intervention	15.2	13.3	21.7	29.8	9.5
Only available Assessment for population	12.8	1.7	13.0	2.1	9.5
Quick to score	12.8	11.7	21.7	**72.3**	**42.9**
Good for selecting class strategies	8.8	3.3	8.7	6.4	21.5
Assessment includes reading/writing	6.0	8.3	4.3	0.0	2.4
Assessment includes social abilities	4.0	5.0	0.0	1.1	0.0
Recently developed Norms	1.6	40.0	4.3	0.0	**42.9**
Inexpensive	0.0	0.0	0.0	17.0	11.9
Other Reason	3.2	0.0	2.2	3.2	0.0

Standardized measures have set guidelines for administration and scoring. Regular use was defined if the measure was used with half or more than half of the last 40 children. The sample size (n) is different for each measure as not all SLPs used all measures regularly. This table shows the percentage of SLPs who identified each reason as a main reason for use. SLPs could rank up to three main reasons for each measure. A reason was taken to be a "main reason" if it was one of the (up to three) reasons selected by an SLP. On average, each SLP selected 2.9 reasons for each measure. Bold font: indicates the three most frequently selected reasons for each measure.

a focus for intervention, thus reducing the impact of intervention for school-aged children [11, 62, 66]. In addition, these measures may not be sufficiently sensitive to measuring changes in language ability over time [10]. As such, use of diagnostic measures such as the CELF core language and language index subtests for purposes of selecting intervention or detecting change does not align well with evidence-based practice recommendations [10, 66]. Intervention goals and targets are more appropriately identified and monitored using assessment tasks that are representative of the child's everyday communication and functional for an individual child's specific communication needs [11, 14].

In this survey, 80% of SLPs who reported regularly using the RAPT used this measure for screening purposes. As the RAPT does not have evidence of diagnostic accuracy [20, 67], it is not identified as being appropriate for screening language abilities. This measure contains only 10 short answer questions for children, which may not be a sufficient number of test items for accurate screening. Furthermore, at the time of this survey, the RAPT had normative data collected 30 years ago [20], thus making the norms outdated. Unlike the CELF measures, quality of psychometric properties was not frequently identified as a main reason for choosing the RAPT by SLPs in this survey. This suggests that SLPs

clinical decision-making may be different depending on the particular measure chosen for use. 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. Instead, SLPs reported selecting the RAPT due to this measure being quick to administer and score. While it is important that consideration be given to the cost of professional time, it is important that this be balanced with evidence-based practice recommendations. Given the limitations of the RAPT, there is an identified need for change with regards to frequent use of the normative data from the RAPT for screening purposes.

In relation to versions of the SPAT, the most frequently selected purpose was "selecting intervention." SLPs also indicated that the SPAT was "good for selecting intervention goals." Given that the SPAT assesses a range of phonemic awareness abilities that are important for word reading and spelling, the SPAT is likely useful for informing phonemic awareness and literacy interventions. However, "diagnosis" and "screening" were each selected as main purposes of the SPAT by approximately 40% of SLPs, despite the SPAT not having evidence of diagnostic accuracy. Furthermore, over 42% of SLPs indicated that "recently developed norms" were a main reason that they chose the SPAT for use. However, at the time of this survey, the most recent version of the SPAT had norms that were over 15 years old and not considered recent. Given these findings, it is suggested that Australian SLPs give greater consideration toward the reasons that they choose specific standardized measures for use.

Clinical Implications

Findings from this survey suggest that SLPs may use standardized measures to assess domains that the measures are not designed for and for purposes for which measures are not well suited. Although the focus of this study was standardized measures used regularly by SLPs in Australia, SLPs internationally are encouraged to critique their own use of standardized language measures in relation to evidence-based practice. Using standardized measures inappropriately may have significant clinical consequences as the foundation of decisions made from assessment data may be compromised. It is acknowledged that SLPs combine data from standardized measures with additional information from non-standardized procedures when undertaking language assessment with school-aged children; however, the use of non-standardized procedures does not counteract for improper use of standardized measures. To assist in building SLP knowledge regarding purpose of assessment, it is recommended that both undergraduate training and post-graduate continuing professional development for SLPs place greater emphasis on the specific domains and purposes for which different language measures are suitable.

Findings from this survey also identified that although only a small number of standardized measures are used regularly, a vast array of standardized measures are available for assessing the language abilities of school-aged children. Having such a large array of options may be over-whelming and lead to "choice-overload," thus making it difficult for SLPs to make sound decisions about which measures to use [68]. The development of resources such as clinical practice guidelines or decision-making aids may be needed to assist SLPs to make evidence-based decisions when choosing standardized measures to assess the language abilities of school-aged children.

Limitations and Future Directions

This study surveyed a varied sample of SLPs from different geographical locations and with different backgrounds; however, as with any survey, it cannot be ensured that findings from this survey are representative of all SLPs. Despite the provision of agreed-upon definitions within the survey for describing language assessment domains and purposes [3], it is possible that some survey participants may not have accurately identified all the standardized measures they used. There was also a proportion of standardized measures that could not be identified from participant descriptions and were thus not able to be included. The survey questions needed to be accompanied by examples to support consistent application of terms; however, it is acknowledged that participants may have been primed by the examples provided. It is hoped that future surveys could also utilize the same consistent terminology for language assessment and SLP practice to facilitate comparison across studies and replication of findings.

The purpose of this study was to examine SLP's use of standardized language measures in relation to evidence-based practice. Therefore, this study focused on the domains, purposes, and reasons for which the most regularly used standardized measures were used. Due to the need to keep the survey length manageable for participants, data were not collected in this survey on the purposes, domains, and reasons for which SLPs use standardized measures that are otherwise used infrequently. It was also beyond the scope of this study to examine SLP use of non-standardized assessment procedures. Therefore, findings in this study relate to SLPs' use of standardized measures as opposed to an investigation of SLPs overall assessment process. Further qualitative

studies, for example, semi-structured interviews, are needed to explore the decisions SLPs make when selecting irregularly used standardized language measures or non-standardized procedures for use.

Similarly, this study did not examine the decision-making processes SLPs employ when using individual subtests within a standardized measure or when analyzing data from different assessments. Future research using qualitative methods may assist in understanding the decisions SLPs make when combining data from a range of assessments, including individual subtests, to make judgments on the language abilities of school-aged children [6, 7]. An analysis of the factors (e.g., service context, SLP demographics, geographical location) that influence SLP assessment practice more broadly has been previously reported in a separate publication [12]. It was identified that SLP assessment practice may be influenced by service agency, SLP years of experience, and geographical location [12]. Future research should also investigate how SLP assessment practice is influenced by children's linguistic profiles. Case studies with children with varied profiles and suspected diagnoses may be helpful in further understanding the factors that influence SLP language assessment practices.

Conclusion

This study investigated the standardized measures used by Australian SLPs when assessing the language abilities of school-aged children. As a group, SLPs listed many standardized measures; however, only a small number of these measures were used regularly by SLPs. SLPs reported using standardized measures to target domains that measures are not ideally designed to target or purposes that measures are not ideally suited to measuring. In addition, there is an identified need to improve SLP decision-making in relation to reasons why measures are chosen for use. Overall, these findings identify the need for SLPs to reflect on their clinical reasoning in relation to evidence-based practice recommendations when selecting standardized measures to assess the language abilities of school-aged children.

Acknowledgment

The authors would like to thank the Australian Government Research Training Program Scholarship that supported this research.

Statement of Ethics

Ethical approval for this survey was provided by the Curtin University Ethics Committee (Approval number: HRE2017-0659). Participants were presented with information on the study at the start of the survey and were required to answer "yes" to the first survey question asking for consent to participate before proceeding to the survey questions. Participants who did not provide consent were not able to access the survey. All participants in this study were over the age of 21 years.

Conflict of Interest Statement

There are no conflicts of interest in relation to this paper.

Funding Sources

There are no funding sources in relation to this paper.

Author Contributions

Deborah Denman: conceptualization, methodology, formal analysis, investigation, writing – original draft preparation, and project administration. Reinie Cordier: conceptualization, methodology, formal analysis, writing – review and editing, and supervision. Jae-Hyun Kim, Natalie Munro, and Renee Speyer: conceptualization, methodology, writing – review and editing, supervision.

Data Availability Statement

All data generated or analyzed during this study are included in this article. Further inquiries can be directed to the corresponding author.

References

1 Bishop DVM, Snowling MJ, Thompson PA, Greenhalgh T; The CATALISE-2 consortium. Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: terminology. J Child Psychol Psychiatry. 2017;58(10):1068–80.

2 Norbury CF, Gooch D, Wray C, Baird G, Charman T, Simonoff E, et al. The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. J Child Psychol Psychiatry. 2016;57(11):1247–57.

3 Denman D, Kim JH, Munro N, Speyer R, Cordier R. Describing language assessments for school-aged children: a Delphi study. Int J Speech Lang Pathol. 2019;21(6):602–12.

4 Kaderavek JN. Assessment of Language disorders. Language disorders in children: fundamental concepts of assessment and intervention. US: Pearson; 2015. p. 45–95.

5 Fulcher-Rood K, Castilla-Earls A, Higginbotham J. Diagnostic decisions in child language assessment: findings from a case review assessment task. Lang Speech Hear Serv Sch. 2019;50(3):385–98.
6 Fulcher-Rood K, Castilla-Earls AP, Higginbotham J. School-based speech-language pathologists' perspectives on diagnostic decision making. Am J Speech Lang Pathol. 2018;27(2):796–812.
7 Selin CM, Rice ML, Girolamo T, Wang CJ. Speech-language pathologists' clinical decision making for children with specific language impairment. Lang Speech Hear Serv Sch. 2019;50(2):283–307.
8 Spaulding TJ, Swartwout Szulga M, Figueroa C. Using norm-referenced tests to determine severity of language impairment in children: disconnect between US policy makers and test developers. Lang Speech Hear Serv Sch. 2012;43(2):176–90.
9 Betz SK, Eickhoff JR, Sullivan SF. Factors influencing the selection of standardized tests for the diagnosis of Specific Language Impairment. Lang Speech Hear Serv Sch. 2013;44(2):133–46.
10 Dockrell JE, Marshall CR. Measurement issues: assessing language skills in young children. Child Adolesc Ment Health. 2015;20(2):116–25.
11 Ogiela DA, Montzka JL. Norm-referenced language test selection practices for elementary school children with suspected Developmental Language Disorder. Lang Speech Hear Serv Sch. 2021;52(1):288–303.
12 Denman D, Cordier R, Kim J-H, Munro N & Speyer R. What Influences Speech-Language Pathologists' Use of Different Types of Language Assessments for Elementary School–Age Children? Lang Speech Hear Serv Sch. 2021 52(3), 776–793.
13 Bishop DVM, Snowling MJ, Thompson PA, Greenhalgh T; CATALISE consortium. CATALISE: a multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. PLoS One. 2016;11(7):e0158753.
14 Kover ST, Davidson MM, Sindberg HA, Ellis Weismer S. Use of the ADOS for assessing spontaneous expressive language in young children with ASD: a comparison of sampling contexts. J Speech Lang Hear Res. 2014;57(6):2221–33.
15 Volden J, Coolican J, Garon N, White J, Bryson S. Brief report: pragmatic language in autism spectrum disorder: relationships to measures of ability and disability. J Autism Dev Disord. 2009;39(2):388–93.
16 Wade DT. Assessment, measurement and data collection tools. Clin Rehabil. 2004;18(3):233–7.
17 Glover A, McCormack J, Smith-Tamaray M. Collaboration between teachers and speech and language therapists: services for primary school children with speech, language and communication needs. Child Lang Teach Ther. 2015;31(3):363–82.
18 Shahmahmood TM, Jalaie S, Soleymani Z, Haresabadi F, Nemati P. A systematic review on diagnostic procedures for specific language impairment: the sensitivity and specificity issues. J Res Med Sci. 2016. 21. 67.
19 Caesar LG, Kohler PD. Tools clinicians use: a survey of language assessment procedures used by school-based speech-pathologists. Commun Disord Q. 2009;30(4):226–36.
20 Renfrew C. Renfrew action picture test. Camberwell, Australia: ACER; 2010.
21 Letts C, Edwards S, Schaefer B, Sinka I. The new Reynell Developmental Language scales: descriptive account and illustrative case study. Child Lang Teach Ther. 2014;30(1):103–16.
22 Roulstone S, Marshall JE, Powell GG, Goldbart J, Wren YE, Coad J, et al. Evidence-based intervention for preschool children with primary speech and language impairments: child Talk – an exploratory mixed-methods study. Programme Grants Appl Res. 2015;3(5):1–408.
23 Wiig EH, Secord WA, Semel E. Clinical evaluation of Language Fundamentals: Preschool–2nd Edition. 2nd ed. Bloomington (MN): Pearson; 2004.
24 Wiig EH, Semel E, Secord WA. Clinical evaluation of Language Fundamentals: 4th Edition. 5 ed. Bloomington (MN): Pearson; 2004.
25 Wiig EH, Semel E, Secord WA. Clinical evaluation of Language Fundamentals. 5th Edition. Bloomington (MN): Pearson; 2013.
26 Sharma A, Minh Duc NT, Luu Lam Thang T, Nam NH, Ng SJ, Abbas KS, et al. A consensus-based checklist for reporting of survey studies (CROSS). J Gen Intern Med. 2021;36(10):3179–87.
27 Qualtrics. Qualtrics March 2018 edition. Provo, Utah, USA: Qualtrics; 2005.
28 Blais JG, Grondin J. The influence of labels associated with anchor points of likert-type response scales in survey questionnaires. J Appl Meas. 2011;12(4):370–86.
29 Andrews D, Nonnecke B, Preece J. Electronic survey methodology: a case study in reaching hard to involve Internet Users. Int J Hum Comput Interact. 2003;16(2):185–210.
30 IBM Corp. IBM SPSS Statistics for windows version 20.0. Released 2011. NY: IBM Corp: Armonk.
31 Gillam RB, Pearson NA. Test of Narrative Language: Second Edition. Austin (TX): PRO-Ed Inc; 2017.
32 Neilsen R. Sutherland phonological awareness test: revised. Camberwell, Australia: ACER; 2003.
33 Zimmerman IL, Steiner VG, Pond RE, Preschool Language scales: 5th ed. 5th ed. Bloomington (MN): Pearson; 2011.
34 Gillam RB, Pearson NA. Test of Narrative Language (TNL). Austin (TX): PRO-Ed Inc; 2004.
35 Bowers L, Huisingh R, LoGiudice C. Test of problem solving 3rd ed (TOPS-3). East Moline (IL): LinguiSystems, Inc; 2005.
36 Snowling MJ, et al. York assessment of reading comprehension (YARC). London, England: GL Assessment; 2009.
37 Renfrew C. Renfrew bus story. Camberwell, Australia: ACER; 2010.
38 Bishop DVM. Children's communication checklist: Second Edition. London, England: Psycholgical Corporation; 2003.
39 Neale MD. Neale analysis of reading: Third Edition. Camberwell, Australia: ACER; 1999.
40 Dunn LM, Dunn DM. Peabody picture vocabulary test: fourth edition. 4th ed. Bloomington (MN): Pearson; 2007.
41 Edwards S, Letts C, Sinka I. New Reynell Developmental Language scales. London, England: GL Assessment; 2011.
42 Wagner R, et al. Comprehensive test of phonological processing: Second Edition. Austin (TX): ProEd; 1999.
43 Dodd B. The university of queensland dept. Of speech Pathology and audiology, Queensland University Inventory of Literacy (QUIL). Brisbane, Australia: The University of Queensland Dept. of Speech Pathology & Audiology; 1996.
44 Semel E, Wiig E, Secord W. Clinical evaluation of Language Fundamentals: fourth edition screening test (CELF-4 screening). Bloomington (MN): Pearson; 2006.
45 Rowland C, Fried-Oken M. Communication matrix. 2004. [cited 2018 Oct 2018]. Available from: https://communicationmatrix.org/.
46 Newcomer PL, Hammill DD. Test of Language development: intermediate–4th ed. 4th ed. Austin (TX), U.S.: Pro-Ed; 2008.
47 Carrow-Woolfolk E. Comprehensive assessment of spoken language: Second Edition. Torrance (CA): WPS; 2017.
48 Bishop DVM. Expression, Reception and Recall of narrative instrument. US: Pearson; 2004.
49 Carrow-Woolfolk E. Oral and written language Scales: Second Edition. 2nd ed. Bloomington (MN): Pearson; 2011.
50 Hammill DD, Newcomer PL. Test of oral language development: primary–4th Edition. Austin (TX): Pro-Ed; 2008.
51 Health Commission of New South Wales. The Bureau auditory comprehension test. Sydney, Australia: Health Commission of New South Wales; 1990.
52 Leitao S, Allan L. Peter and the cat narrative assessment. Black Sheep Press Ltd; 2003.
53 Blank M, Rose SA, Berlin LJ. Preschool language assessment instrument (PLAI-2). Austin (TX): ProEd; 2003.
54 Phelps-Terasaki D, Phelps-Gunn T. Test of pragmatic language: 2nd ed (TOPL-2). Torrance (CA): WPS; 2007.
55 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.
56 Westwood P. Spelling: approaches to teaching and assessment. Australian Council for Educational Research; 2005.

57 Watson RM, Pennington L. Assessment and management of the communication difficulties of children with cerebral palsy: a UK survey of SLT practice. Int J Lang Commun Disord. 2015;50(2):241–59.
58 Dethorne LS, Johnson BW, Loeb JW. A closer look at MLU: what does it really measure? Clin Linguist Phon. 2005;19(8):635–48.
59 Ukrainetz TA, Blomquist C. The criterion validity of four vocabulary tests compared with a language sample. Child Lang Teach Ther. 2002;18(1):59–78.
60 Caesar LG, Kohler PD. The state of school-based bilingual assessment: actual practice versus recommended guidelines. Lang Speech Hear Serv Sch. 2007;38(3):190–200.
61 Hunt E, Nang C, Meldrum S, Armstrong E. Can dynamic assessment identify language disorder in multilingual children? Clinical applications from a systematic review. Lang Speech Hear Serv Sch. 2022;53(2):598–625.
62 Trembath D, Westerveld M, Shellshear L. Assessing spoken language outcomes in children with ASD: a systematic review. Curr Dev Disord Rep. 2016;3(1):33–45.
63 Lennox M, Westerveld MF, Trembath D. Should we use sentence-or text-level tasks to measure oral language proficiency in year-one students following whole-class intervention? Folia Phoniatr Logop. 2018;69(4):169–79.
64 Bishop DVM, McDonald D. Identifying language impairment in children: combining language test scores with parental report. Int J Lang Commun Disord. 2009;44(5):600–15.
65 Denman D, Speyer R, Munro N, Pearce WM, Chen YW, Cordier R. Psychometric properties of language assessments for children aged 4–12 years: a systematic review. Front Psychol. 2017;8:1515–28.
66 Ebert KD, Scott CM. Relationships between narrative language samples and norm-referenced test scores in language assessments of school-age children. Lang Speech Hear Serv Sch. 2014;45(4):337–50.
67 Glover TA, Albers CA. Considerations for evaluating universal screening assessments. J Sch Psychol. 2007;45(2):117–35.
68 McCabe PJ. 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. Int J Speech Lang Pathol. 2018;20(3):300–9.