

## Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature \_\_\_\_\_

Date \_\_\_\_\_

FACTORS INFLUENCING STANDARDIZED TEST SELECTION  
FOR CHILDREN PRESENTING WITH LANGUAGE DIFFICULTIES

by

Jennifer L. Montzka

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in the Department of Communication Sciences and Disorders

Idaho State University

Summer 2015

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

## Committee Approval

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Jennifer L. Montzka find it satisfactory and recommend that it be accepted.

---

Diane A. Ogiela, Ph.D., CCC-SLP  
Major Advisor

---

Jeanne Johnson, Ph.D., CCC-SLP  
Committee Member

---

Elizabeth Doughty Horn, Ph.D.  
Graduate Faculty Representative

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Copyright 2015 Jennifer L. Montzka

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

# Idaho State UNIVERSITY

Office for Research Integrity  
921 South 8th Avenue, Stop 8046 • Pocatello, Idaho 83209-8046

November 13, 2014

Diane Ogiela  
1311 E Central Ave  
CSED  
Meridian, ID 83642

RE: Your application dated 11/11/2014 regarding study number 4195: Factors Influencing Assessment Instrument Selection for Children Presenting With Language Difficulties

Dear Dr. Ogiela:

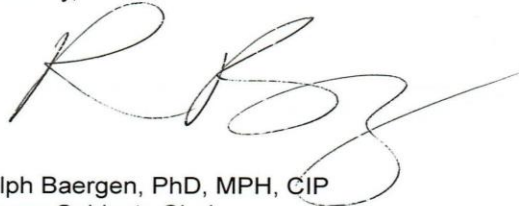
I agree that this study qualifies as exempt from review under the following guideline: 1. Research on educational practices in educational settings. This letter is your approval, please, keep this document in a safe place.

Notify the HSC of any adverse events. Serious, unexpected adverse events must be reported in writing within 10 business days.

You are granted permission to conduct your study effective immediately. The study is not subject to renewal.

Please note that any changes to the study as approved must be promptly reported and approved. Some changes may be approved by expedited review; others require full board review. Contact Tom Bailey (208-282-2179; fax 208-282-4723; email: humsubj@isu.edu) if you have any questions or require further information.

Sincerely,



Ralph Baergen, PhD, MPH, CIP  
Human Subjects Chair

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### Acknowledgements

Thanks to the entire 2015 graduate class for their support, enthusiasm and kindness.

Additional thanks goes to Melissa DeLamere, Allison Early, Nicole Quirk, Laura Rogers, Erika Simoneau, Michelle Wallior, Sara Wenig, and Heather Ybarlucea for assistance in testing, proof-reading, and providing feedback on the survey.

Thanks also go to Mary Van Donsel, Darcey Kennedy, and Jody O'Donnell for additional feedback on the survey.

Special thanks go to Dr. Kristina Blaiser for feedback and recommendations on the survey and prospectus project.

Thanks to Dr. Jeanne Johnson, for providing guidance, support, and feedback above and beyond the call of duty throughout the development, prospectus, and completion of this project.

Additional thanks to Dr. Elizabeth Horn for perspective, patience, and thoughtful questions and comments.

Thanks are also due to my fellow minions in the Child Language Lab: Elin Kuzmak, for recruiting research and support; Annabel Armstrong for research support, and Jeanie Chapin and Marissa Dey, for survey testing and research support.

A thank you is also due to my dear friend Katherine Huebschmann and my sister, Kelsey Montzka-Boettiger, for proof-reading, trouble-shooting, and moral support.

Special thank you to my parents, who have not only provided material and emotional support throughout this project, but also raised me with a curious mind and a desire to find answers for all of my many questions.

Finally, I want to thank Dr. Diane Ogiela for bringing me into her lab, convincing me that doing a thesis really was a good idea, and then providing me with support, guidance, constructive criticism and general brilliance throughout the process.

I am so grateful for all she has invested in this project and in me, both as a student and as a young, learning clinician.

Thank you, Dr. O.

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

## TABLE OF CONTENTS

List of Figures .....	ix
List of Tables.....	x
Abstract .....	xi
Background .....	4
Features of Specific Language Impairment in Children .....	4
Psychometric Qualities of Tests .....	9
Evidence-Based Practice.....	15
Assessment Challenges for School-Based SLPs.....	17
Methods.....	20
Procedure .....	20
Questionnaire .....	22
General Assessment Procedures .....	22
Standardized Test Preferences. ....	23
Factors Impacting Test Selection .....	24
Participants.....	26
Licensure.....	27
Region .....	28
Years of Experience .....	28
Results.....	29
General Assessment Procedures .....	29
Initial Test Selection.....	30
Additional Test Selection .....	32

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Factors Impacting Test Selection .....	37
Most and Least Important Factors .....	42
Sources of Information Regarding Test Psychometric Quality .....	44
Discussion .....	49
Test Selection .....	49
Factors Impacting Test Selection .....	54
Years of Experience and Factors Influencing Test Selection .....	55
Source of Psychometric Information and Factors Influencing Test Selection .....	56
Best Practice Recommendations and Factors Impacting Test Selection .....	58
Limitations .....	59
Conclusions and Future Directions .....	60
References .....	62
Appendices .....	72
Appendix A. Questionnaire	
Selection includes initial, elementary grade, and demographics portions .....	72
Appendix B. List of Tests Provided to Participants .....	83
Appendix C. Additional Assessments Recommended .....	87
Appendix D. States Listed by Regions .....	89
Appendix E. Participant-generated Most Important and Least Important <i>Factors</i> .....	90



# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

## LIST OF FIGURES

Figure 1. Regional distribution of participants .....	28
Figure 2. Years of participant experience.....	28
Figure 3. Participants' selections of initial tests.....	30
Figure 4. Domains assessed by participants' initial test selections.....	31
Figure 5. Additional test selection by percentage of participants. ....	33
Figure 6. Percentage of additional test domains selected out of total number of additional tests selected.....	34
Figure 7. Domains of additional tests by purposes for using additional tests.....	36
Figure 8. The impact of factors on test selection by percentage of respondents .....	38
Figure 9. Relationships between years of experience and test selection factor ratings .....	40
Figure 10. Participants' selections of most and least important factors.....	42
Figure 11. Self-identified most important test selection factor.....	43
Figure 12. Differences in initial test selection by participants who identify test- characteristic factors or other characteristics as most important .....	44
Figure 13. Sources of information participants use to gather information about test psychometric quality.....	45
Figure 14. Relationships between sources of psychometric information and test selection factor ratings .....	46
Figure B1. Map provided to questionnaire participants for demographics questions.....	89

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

## LIST OF TABLES

Table 1. Significant Spearman's $\rho$ Correlations of Test Selection Factors .....	39
Table 2. Multiple Comparisons: Years of Experience & Test Selection Factor Ratings....	41
Table 3. Multiple Comparisons: Source of Psychometric Information & Test Selection Factor Ratings .....	47
Table B. <b>Error! Reference source not found.</b> .....	83
Table C. Additional Assessments Recommended.....	87
Table D. States Listed by Regions .....	89
Table E1. Self-generated Most Important Factor.....	90
Table E2. Self-generated Least Important Factor .....	92

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### Abstract

The evaluation of language problems in school-age children is a primary responsibility of school speech-language pathologists (SLPs). Standardized norm-referenced tests are a common part of this multi-faceted process. There are many tests from which to choose, making the process of selecting appropriate tests a challenge. This study explored the assessment practices of school-based SLPs, examining possible influences on the selection of standardized tests for children presenting with language deficits. Results from the study revealed that situational factors such as availability and personal familiarity were rated as having the greatest impact on test selection decisions and test-quality factors such as reliability, validity and diagnostic accuracy were rated as having less impact. Best practice recommends that factors related to test-quality, such as psychometric features, be a driving force in test selection, but a number of situational factors make the examination and evaluation of these test-quality features challenging for school-based SLPs.

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

## Factors Influencing Assessment Instrument Selection for Children Presenting With Language Difficulties

The identification and treatment of children with language deficits is a primary responsibility of school-based speech-language pathologists (SLPs) according to the American Speech-Language-Hearing Association (ASHA, 2010). Results from ASHA's most recent survey of school-based SLPs indicate that 69% serve children with language difficulties other than deficits in pragmatics and/or social communication (ASHA, 2012). Children with language difficulties often present with a variable profile of social, behavioral, and communication deficits, but because their difficulties are caused by an underlying language disorder, the SLP is the principal professional responsible for making a diagnosis (Betz, Eickhoff, & Sullivan, 2013).

Responsibility for the diagnosis of language deficits requires SLPs to accurately identify and assess children who present with language deficits. The use of standardized norm-referenced assessment instruments (tests) is a frequent, and sometimes required, element of the process of language assessment (Betz et al., 2013; Caesar & Kohler, 2009; Wilson, Blackmon, Hall, & Elcholtz, 1991). However, past research has demonstrated that the many language tests available to clinicians vary widely in psychometric quality (Betz et al., 2013; Friberg, 2010; McCauley & Swisher, 1984; Plante & Vance, 1994) and level of diagnostic accuracy (Spaulding, Plante, & Farinella, 2006). This makes the selection of an appropriate test an even more challenging task for clinicians.

A test that is not appropriate for a client may contribute to an inaccurate diagnosis, which places the child at a higher risk for academic, educational, and social difficulties due to his or her language difficulties (Hoffman, Loeb, Brandel, & Gillam,

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

2011). With this in mind, it is clear that best practice in speech-language pathology describes accurate assessment is an important responsibility for SLPs(ASHA, 2004; ASHA, 2010).

The process of selecting a test that is appropriate for a client requires a clinician to evaluate multiple factors and use clinical judgment to determine the best possible course of action. When language tests are appropriately selected, they can provide a great deal of valuable information to the SLP, leading to more accurate diagnoses and more effective treatment. If tests are not selected appropriately, the information they provide may be less valuable or misleading. Therefore, it is critical to identify and analyze the factors that influence SLPs' decision-making processes when selecting tests to identify any possible barriers to effective practice.

The purpose of this paper is to evaluate and discuss the results from a national survey of school-based SLPs regarding which factors impact their selection of standardized language tests for children presenting with language difficulties. The results from SLPs who have caseloads that include at least 30% elementary (first through fifth grade) students are investigated here. This study examined which factors school-based SLPs consider more or less important when selecting standardized language tests, and explored the relationships between factors, test selection choices, region, and years of experience.

In addition to understanding which standardized tests are most frequently administered as primary tests, and which are most frequently administered in addition to a primary language test, the most important questions explored are those related to the factors that influence test selection. The primary questions addressed here include: (a)

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

which factors are most and least important to school SLPs, (b) how are factors related to each other, on a continuum of strongly affecting to not affecting selection decisions, (c) what is the relationship between the source of information about test psychometric quality (i.e. informal reports, personal judgment, reviews of tests published in peer-reviewed journals, the tests' examiner's manual, or the publisher's catalog description) and test selection factors, (d) and finally (d) do SLPs' years of experience and region relate to the most and/or least important factors in their decision-making process.

Information provided by this research has potentially important implications for clinical practice in a school-based setting. The first aim of this research is to provide a description of current practice patterns, enabling comparisons between current practice and best practice recommendations. Next, it offers an exploration of which factors of the test selection process make particular tests more preferable to school-based SLPs. Finally, by examining which factors are more important and which factors are less important to school-based SLPs, it is possible to examine the information available to SLPs and how this information may impact their selection process.

### Background

#### Features of Specific Language Impairment in Children

When a child presents with language difficulties without other developmental disabilities, neurological complications, or sensory deficits, the language impairment may be identified by the diagnostic label of *specific language impairment* (SLI; See Leonard, 2014 for a review). As found by Tomblin et al. (1997), SLI occurs in approximately 7% of the kindergarten population. Tomblin's epidemiological cross-sectional study of over 7,000 kindergarten students is a level 2+ study, as it is a well-designed, prospective, non-experimental study (Harbour & Miller, 2001). The researchers recruited a large number of diverse participants from rural, urban, and suburban populations, resulting in a representative sample that is considered likely to generalize to broader populations (Tomblin et al., 1997). Despite the high prevalence reported by Tomblin's study, SLI is not widely-recognized by the general public and educators. This is true even though SLI appears to be more prevalent among children than more commonly recognized disorders such as autism spectrum disorder, which occurs at a rate of approximately 1.13% (Center for Disease Control and Prevention, 2008), and attention deficit hyperactivity disorder, which occurs at a rate of approximately 5% (American Psychiatric Association, 2013). SLI may be under-identified due to the 'hidden' nature of the disorder, as difficulties in understanding and expressing language may not be noticed until they result in behavioral challenges or academic deficits in reading and writing (Bishop & Clark, 2012; Rice, 2014).

It is important to note that the term *SLI* is not universally used or accepted throughout the profession. There is a lack of consensus among professionals regarding

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

the best diagnostic label to use, resulting in difficulty providing services to these children in clinical settings (Dockrell & Lindsay, 2014). Some professionals contest the use of the term ‘specific’ as a part of the diagnostic label. They question the validity of a unique clinical diagnostic category (as opposed to “language impairment”), the necessity for adequate clinical intervention, and the reality of a language deficit truly occurring in isolation (Reilly et al., 2014; Rutter, 2014). Critics of the term *specific language impairment* often recommend the use of other terms such as *language impairment*, *primary language impairment*, *developmental language disorder*, or *language learning impairment* (Bishop, 2014). These terms are not without their difficulties, however. As discussed in Grist and Hartshorne (2014), *language impairment* does not distinguish between those who have this impairment in addition to other disorders, such as Autism Spectrum Disorder, and the term can be used to describe acquired language disorders in adults. Parsons, Jordan, & Branagan (2014) point out that primary language impairment, when abbreviated as *PLI*, could be mistaken for *pragmatic language impairment* or the word *primary* would be incorrectly associate the disorder with only primary grades. Bishop (2014) notes that changing the name drastically, such as re-naming the disorder *language learning impairment*, could result in past research on the disorder being harder to find or becoming disconnected from new research. Leonard (2014b) suggests that some opposition to using SLI as a diagnostic term may come from adhering to outdated definitions of the disorder from older research and expresses concern that broader label such as *language impairment* could disrupt interdisciplinary communication and impact children who are in need of language services. The limited acceptance of a single term for the disorder is thought be a contributing factor to the limited degree of research funding



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

directed to studying SLI and limited public recognition and understanding of the disorder (Bishop, 2010; Grist & Hartshorne, 2014). Detailed discussion of the appropriate term to use is beyond the scope of the present paper. Given that the vast majority of the research on this disorder has been conducted using the label SLI, that term will be used here.

The language of those children identified as having SLI may present as deficits in several different areas of language. Because of the heterogeneity that exists among children with SLI, a great deal of research has been conducted seeking to identify the patterns of impairment in this population. This research has indicated several key language characteristics that are often exhibited by children with SLI.

Children with SLI often begin saying their first words later in than their same-age typical peers and their overall lexical development is often slower (Leonard, 2014a). A meta-analysis conducted by Kan and Windsor (2010) concluded that children with SLI do not perform as well as children with typical language on novel word-learning tasks. McGregor, Oleson, Bahnsen, and Duff (2013) found that children with SLI often demonstrate a less diverse vocabulary, a delay in combining words, and greater difficulties with expressive language. Numerous researchers have also found that young children with SLI demonstrate marked weakness in grammatical morphology, particularly as it relates to tense marking (Bedore & Leonard, 1998; Charest & Leonard, 2004; Conti-Ramsden, 2003; Leonard, 2014a; Rice & Wexler, 1996; Rice, Wexler, Marquis, & Hershberger, 2000). Children with SLI also exhibit slower syntactic development, as preschool children with SLI use fewer complex syntactic structures than typical peers (Schuele & Dykes, 2005; Arndt & Schuele, 2013). In addition, children with

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

SLI have a relatively high rate of concomitant speech-sound disorders (Shriberg, Tomblin, & McSweeney, 1999).

As children mature, however, the disorder presents differently, often characterized by difficulty with syntactic complexity (Leonard, 1995; Nippold, Mansfield, Billow, & Tomblin, 2008; Schuele & Dykes, 2005; Scott & Stokes, 1995). Specifically, school-age children with SLI often use fewer clauses and less subordination than their same-age peers (Scott & Balthazar, 2010). Errors of tense & agreement, while no longer occurring frequently in spoken language, often persist in written language. (Bishop & Clarkson, 2003; Windsor, Scott, & Street, 2000). Adolescents with SLI also demonstrate less narrative complexity (Koutsoftas & Gray, 2012) and shorter mean length of spoken utterances (Nippold et al., 2009).

The majority of the studies that evaluate the characteristics of SLI are cohort studies with some risks of confounding bias, usually attributable to sampling bias in participant selection or participant attrition rates across time. However, there has also been replication of many of the findings with regard to the characteristics of SLI. According to the system for grading recommendations in evidence based guidelines by Harbour and Miller (2001), the levels of evidence ratings fall between a 2+ (well-conducted cohort studies with a low risk of confounding bias) and a 2- (cohort studies with a high risk of confounding bias) except for Kan and Windsor (2010)'s meta-analysis, which demonstrates a 1+ (well-conducted meta-analysis with a low risk of bias) level of evidence.

Behavioral, social, and emotional difficulties (BSED) can be associated with SLI as well. A level 2+ (well-conducted cohort studies with a low risk of confounding bias)

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

study of 234 to 103 children identified as having SLI participated in reading, oral language, and BSED evaluations. This study indicated that as children with SLI transition from childhood into adolescence, hyperactivity and conduct difficulties tend to decrease while social difficulties increase, likely due to an increased emphasis on language skills in peer interactions during the adolescence (St Clair, Pickles, Durkin, & Conti-Ramsden, 2011). Stringer and Lozano's 2007 study, rated at a level of evidence of 2- (cohort studies with a high risk of confounding bias), with 21 students (age 8;3 – 13;5) attending a specialized program for children with emotional behavioral disorders indicated that more than half of children in the study also had language impairments and that less than half of these children with language difficulties had been appropriately identified.

Altogether, children with SLI demonstrate difficulty in all domains of language, although for some, semantic and pragmatic skills are somewhat less impacted than morphosyntax (Leonard, 2014). While these patterns of deficit have been associated with SLI, it is important to note that language profiles for individuals with this disorder are variable. Different individuals will demonstrate varied levels of skill or impairment in the different domains of language, although the domains of morphology and syntax commonly demonstrate the greatest impact over the long term. This heterogeneity of the population of children with SLI necessitates using instruments that provide high quality, specific assessment information regarding each child's particular language deficits. However, not all standardized tests have the same level of psychometric quality and diagnostic accuracy, making the process of using these tests in an effective and evidence-based manner even more complex.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### Psychometric Qualities of Tests

High-quality tests have a number of psychometric features including high degrees of *reliability*, *validity*, and *diagnostic accuracy* (Betz et al., 2013; Friberg, 2010; McCauley & Swisher, 1984; Plante & Vance, 1994; Spaulding, Plante, & Farinella, 2006). The measure of *reliability* describes “the consistency with which a test measures a given attribute or behavior” (McCauley & Swisher, 1984, p. 35). This consistency can be measured by evaluating *inter-rater reliability*, which refers to how consistent the scoring remains when the task or test is administered and scored by different individuals. Another type of reliability, *test-retest reliability*, is defined as “consistency of scores obtained by the same individual when s/he is examined on different occasions, using different sets of equivalent test items, or under testing conditions that vary in some other way” (McGoey, Cowan, Rumrill, & LaVogue, 2010, pp. 105-106).

There are also several different types of *validity*. Validity indicates “whether a test accurately reflects the construct that it was designed to measure” (Dollaghan, 2007, pg. 27). Evidence of validity may come from a number of different sources. This evidence may be based on test content, response processes, internal structure, or the relationship of the construct with a criterion (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014).

Similarly, there are several different measures of *diagnostic accuracy*. *Likelihood ratios* are commonly used in the field of medicine and are considered one of the best measures of diagnostic accuracy (Dollaghan, 2007). This is because they evaluate the probability of both true and false positives in a single measure (the positive likelihood ratio) and the probability of true and false negatives in a single measure (the negative

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

likelihood ratio). Unfortunately, these measures are not commonly calculated for speech-language pathology tests (Betz et al., 2013). Another method of determining diagnostic accuracy is to examine the difference in mean scores between the participants with typical language and participants with language impairment. In order for the group scores to be diagnostically meaningful, the group mean scores should be more than one standard deviation from each other (Spaulding et al., 2006). Spaulding et al., (2006) indicate that tests should, at the very least, provide sensitivity and specificity measures so professionals are able to determine diagnostic accuracy. *Sensitivity* is how accurately a test correctly identifies those who do have a disorder, while *specificity* is how accurately a test correctly excludes those who do not have a disorder (Dollaghan, 2007). Both sensitivity and specificity are considered to be acceptable for diagnostic purposes when they reach a level of 0.8 (Plante & Vance, 1994), that is, when they accurately identify 80% of those with a disorder or exclude 80% of those who do not have a disorder. The accuracy levels of both sensitivity and specificity vary depending on the standard deviation at which they are measured (Spaulding et al., 2006), resulting in varying degrees of diagnostic accuracy depending on the score the child earned.

Language tests have repeatedly been found to vary considerably in their psychometric qualities and diagnostic accuracy. McCauley and Swisher, in their 1984 review of 30 tests, discovered that only four tests met any 4 of the 10 psychometric criteria the researchers had established for high quality tests. Plante and Vance (1994) examined the psychometric features of twenty-one language tests that were normed for use with 4 and 5 year old children. Of the twenty-one tests, only 8 met half or more of McCauley and Swisher's psychometric criteria.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

In the next stage of their study, Plante and Vance (1994) administered the 4 highest rated tests (which had met 6 or more of McCauley and Swisher's psychometric criteria) to 40 preschool children. When the results of the testing were evaluated, the ability of these tests to accurately identify children with the disorder, or the tests' sensitivity, was less than 75% for three of the four tests. The most sensitive of the tests, with a sensitivity of 90%, achieved this at a cutoff score of -3.25 standard deviations below the mean. Specificity, or the ability to accurately eliminate those children who did not have the disorder) was 70% or greater for all three tests, but the standard deviation cutoffs to achieve these diagnostic levels varied considerably, ranging from -3.25 to 0.06 standard deviations. This study suggests that even tests that have comparatively strong psychometric characteristics can still demonstrate weak diagnostic accuracy.

Diagnostic accuracy of language tests was further assessed by Spaulding, Plante, and Farinella (2006). Of the 43 standardized tests reviewed, information regarding sensitivity and specificity, or that data from which to calculate those measures, was available for only 9 of the tests and only 5 reached acceptable accuracy levels of 80% for sensitivity and specificity.

Similarly, Betz et al. (2013) evaluated the quality of the psychometric features of 55 tests and conducted a survey to establish how frequently an SLP would use the listed tests. They then attempted to correlate the psychometric features with frequency of administration. The researchers theorized that psychometric features of reliability and validity would correlate with frequency of test administration. Test reliability was determined using test-retest reliability and standard error of measure. If these values were not available, other measures such as split-half reliability were used. Validity was

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

evaluated by examining a test's criterion-related validity; the relationship of the test to another test with the most similar content was used when comparisons multiple tests were provided.

Betz et al. (2013) postulated that sensitivity and specificity were not likely to correlate with frequency of administration. This is because, in their review of the tests, the researchers found that only 13 of the tests had included information indicating that the tests met acceptable levels of diagnostic accuracy, where 'acceptable' is defined as sensitivity and specificity levels of  $>.80$  (with no standard deviation defined). These levels were reported in either in the tests' manuals or in research articles evaluating the tests. Thirty-three of the tests did not include sensitivity and specificity information at all.

The results of Betz and colleagues' analyses revealed that, as predicted, sensitivity and specificity did not correlate with frequency of test administration. Contrary to expectations, however, validity and reliability also did not correlate with how frequently tests were used. This was partially because the reliability of all tests evaluated was within acceptable limits. Surprisingly, the only aspect of the tests that positively correlated with frequency of administration was recency of publication.

The tests used most frequently were, as predicted, multi-domain language tests or vocabulary evaluations. Unfortunately, of the top 10 most frequently administered tests, only two tests, the *Clinical Evaluation of Language Fundamentals, 4<sup>th</sup> edition (CELF-4*; Semel, Wiig, & Secord, 2003) and the *Clinical Evaluation of Language Fundamentals, Preschool-2 (CELF-P2*; Semel, Wiig, & Secord, 2004), have sensitivity and specificity scores within the acceptable range (i.e.  $\geq 0.80$ ). These results suggest that the

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

psychometric qualities of standardized tests may not be one of the central factors involved in the clinician's process of selecting an appropriate test.

The psychometric qualities of a test are not a clinician's only concern. It is also vital to carefully evaluate the particular skills assessed by the test. Hoffman et al., (2011) examined the concurrent and construct validity of two frequently-used tests of oral language, the *Test of Language Development-Primary, 3<sup>rd</sup> Edition (TOLD:P-3;* Newcomer & Hammill, 1997) and the *Comprehensive Assessment of Spoken Language (CASL;* Carrow-Woolfolk, 1999). While both tests assess oral language, each test approaches it in a different way due to the particular theoretical frameworks of the developers. Unfortunately, these tests appear to have less-than-exemplary construct validity. The authors of the *TOLD-P:3* reported that only one factor of language skill, which they termed general spoken language ability, impacted the results of the core subtests. The authors of the *CASL*, on the other hand, stated that three linguistic factors (described as lexical/semantic skills, syntactic skills, and supralinguistic skills) determined performance on the test. Using principle-component factor analysis, Hoffman et al. (2011) determined that only one factor determined performance on the *CASL* for young children and that the *TOLD-P:3* actually had two significant factors, which the researchers termed "Basic Processing" and "Complex Processing." Although these tests purported to either evaluate specific skills within language (such as lexical/semantic, syntactic, and supralinguistic skills) or evaluate general language skills overall, analysis tests indicate that they may not evaluate language skills in the way the test creators had intended.

This divergence of underlying constructs and actual test function does not only occur on the whole-test level. It may also occur on the individual task level, as tasks



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

designed to evaluate a particular skill may not accurately target only this skill. As Hoffman et al. (2011) discussed, the “Picture Vocabulary” subtest on the *TOLD P:3* (Newcomer & Hammill, 1997), only requires a child to point to a picture matching the word named in the stimulus question. As such, it is a closed task with limited response requirements. However, the “Oral Vocabulary” and “Relational Vocabulary” subtests require the child to describe word meanings or relationships between words in a more open-ended response. As generating a response is a task that has greater linguistic (semantic and otherwise) demands than simply identifying a word by pointing to a picture, a child may perform differently on the “Picture Vocabulary” subtest than on the other semantic subtests because of deficits in other areas of language, not in semantics alone. The differences in task complexity and type of skill(s) addressed in each test may result in very different scores for children with different language skill(s).

Another concerning outcome of the Hoffman et al. (2011) study is the lack of diagnostic agreement between *TOLD-P:3* and the *CASL*. A Pearson correlation found the tests correlated with each other with  $r = .596$ , indicating a strong effect size for behavioral sciences, using Cohen's (1988) guideline (cf. Fergusson, 2009 for different guidelines). Although this may demonstrate a relatively strong relationship between the two tests, a t-test,  $t(215) = -8.1$ ,  $p = .001$ , also revealed significant differences between the two composite test scores when the tests were given to the same children with SLI. The researchers discovered that only 64% of study participants previously identified as having language impairment were identified as such by both tests. Although an effect size of .569 may be considered strong for multifaceted complex behavior such as language from a psychological perspective, from a clinical perspective, it is concerning that the

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

two tests only had 64% diagnostic overlap, especially considering the great deal of emphasis is placed on standardized test scores for determining intervention eligibility.

The selection of an appropriate test is particularly vital to ensure the most effective assessment and course of treatment for children with SLI because, as previously discussed, they frequently demonstrate a great deal of heterogeneity in the expression of their language deficits. The research has demonstrated that selecting a test which appropriately targets a child's areas of deficit is crucial to conducting an accurate assessment. In order to ensure that a test not only targets a child's areas of weakness, but also targets them accurately, measures of quality, such as validity, reliability, and diagnostic accuracy, must also be considered. Using a test that meets these qualifications is an important aspect of evidence-based practice in language assessment, as will be discussed in the next section.

### **Evidence-Based Practice**

Evidence-based practice (EBP) seeks to provide the best possible assessment and treatment for each individual client by integrating three key sources of information: the best and most current clinical research, a clinician's own experience, and the preferences of a specific client (Dollaghan, 2007). A primary feature of EBP involves carefully evaluating assessment and intervention tools to ensure that they are of the highest quality and the most appropriate procedure for a particular client. Dollaghan (2007) details a number of questions that a clinician should use to evaluate an assessment instrument such as a standardized test. These questions evaluate features such as: (a) the use of a reference standard (or means by which to identify disordered versus typically developing skills), (b) the quality of the reference standard, (c) the quality of the research design involved in

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

the development of the instrument, and (d) the strength and accuracy of diagnostic measures known as likelihood ratios. Likelihood ratios, as previously mentioned, take into account the potential for both true and false identification of a disorder (for positive likelihood ratios) and the potential for true and false identification of no disorder (for negative likelihood ratios), making them a very effective way to discuss the diagnostic accuracy of any test.

Many surveys have been conducted to ascertain the degree to which EBP has been implemented in clinical practice (Hoffman, Ireland, Hall-Mills, & Flynn, 2013; Meline & Paradiso, 2003; O'Connor & Pettigrew, 2009; Zipoli & Kennedy, 2005). These surveys consistently report that SLPs struggle to find the time to research issues related to the field, read research articles, and implement current research in practice. The level of evidence provided by the surveys discussed in this section is rated at a level 3, with respondents to the survey numbering between 27 and 2,762.

In a survey study of 2,762 respondents, Hoffman et al. (2013) found that those who were more recent graduates reported having had more direct instruction on the use of EBP; however, this instruction did not appear to have a significant impact on the degree to which they included EBP techniques in their professional practice. Despite the reports of low rates of both EBP activity and education, a majority of the participants in multiple surveys indicated an interest in participating in future EBP education (Hoffman et al., 2013; Meline & Paradiso, 2003; O'Connor & Pettigrew, 2009; Zipoli & Kennedy, 2005).

Huang, Hopkins and Nippold (1997) surveyed 216 SLPs in the state of Oregon. Their results suggested that clinicians with high caseloads were more likely to feel dissatisfaction with the assessment process when they also had less understanding about

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

the psychometric quality of tests. The school-based SLP respondents identified time constraints as their greatest challenge to implementing EBP. This difficulty with time constraints also appeared to express itself in dissatisfaction with standardized tests in general, which was reported more often by clinicians who had larger caseloads. Unfortunately, many of these SLPs with high case loads were also school-based clinicians. The results demonstrated a complex interplay that exists between setting, caseload, and clinician understanding, which ultimately impacts the integration of EBP in clinical assessment methods. Together, the surveys of current practice and EBP suggest that many, if not most, clinicians wish to improve the use of EBP, but that time constraints interfere with their ability to do so.

### **Assessment Challenges for School-Based SLPs**

Because language deficits are the primary feature of SLI, it is critical for speech-language pathologists to accurately assess children presenting with language difficulties. However, the quality of norm-referenced tests varies substantially. Even those tests with high quality may not be appropriate for every client due to the heterogeneous nature of the population of children with SLI. This is further complicated by results from past surveys of practicing professionals, which suggest that SLPs struggle to integrate EBP, including an understanding of the psychometric qualities of tests, into their daily clinical practice. If this is truly the case, clinicians may not be adequately evaluating the tests they select to administer to children presenting with SLI.

Gaining insight into school SLPs' current assessment process, specifically how various factors impact their choice of standardized test to administer, provides an opportunity to investigate current practice patterns. If current practices are not in line

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

with recommended EBP, identifying gaps between policy and practice may be the first step towards more accurate assessment of children with SLI in the future. The factors that may impact the selection of standardized tests that will be investigated in this study include: (a) the time commitment required to administer (not score) the test, (b) the time commitment required to score (not administer) the test, (c) personal familiarity with the test based on past use, (d) specific recommendations from other SLPs, graduate program faculty, or clinical supervisors, (e) reviews of the test published in professional journals, (f) publisher's description of the test, (g) the cost of the test or test materials, (h) the availability of the test, (i) guidelines in a school/facility, (j) guidelines in state and federal laws, (k) psychometric features of the test (such as reliability and validity), and (l) specific measures of diagnostic accuracy (such as mean group differences or sensitivity and specificity) for the test as reported in the examiner's manual or in published research articles. Factors such as reviews of tests in published journals, the psychometric features of the test, and the specific measures of diagnostic accuracy are considered to be *test-quality factors*, while all other factors are considered to be *situational factors*.

This review raises a number of questions related to standardized test selection and use. While past surveys have provided information regarding clinician preferences in standardized, norm-referenced test selection and reviews of tests, particularly in regards to test psychometric qualities, no data has been collected regarding the reasons clinicians select particular tests. This study endeavors to begin answering that question. To ensure that the data from this study could be fairly compared to data from past surveys, questions about standardized language test selection were included in the questionnaire. To begin exploring why clinicians select the tests they do, participants were asked to

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

identify factors that are important or unimportant in their selection process. As assessment is a multifaceted process, participants were asked to rate factors on a spectrum of influence, from strongly affecting SLPs' selection decisions to strongly not affecting selection decisions, and also identify which factors are most and least important. To allow for deeper examination of these factors, participants were also asked where they gathered information regarding test psychometric qualities. Finally, demographic effects were used to evaluate the factors that SLPs indicate are most important in their decision-making process. The following section outlines how the data to answer these questions was collected and analyzed.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### **Methods**

This study is part of a larger project examining the assessment practices of school-based SLPs. A survey investigating practices associated with standardized, norm-referenced test selection was distributed electronically to school-based SLPs working in the United States. To ensure that the questions presented to SLPs remained relevant to their clinical expertise, they were asked early in the survey to identify the primary population with whom they work: preschool and kindergarten students, first through fifth grade students (elementary), or sixth through twelfth grade students (middle and high school). While the larger study collected information regarding practices for all school-based SLPs, the present study provides analysis of only the data regarding assessment practices with elementary students.

### **Procedure**

Prior to data collection, this study was approved by the Idaho State University Institutional Review Board on November 13, 2014, as a study exempt from review under the guidelines of research on educational practices in an education setting.

A questionnaire survey was created in Qualtrics, a web-hosted survey management system. This questionnaire contained a total of 65 questions with three possible survey-completion tracks. Participants were asked to complete as few as 24 questions, if they completed only one track and activated no follow-up questions. Although it was possible for a participant to complete all three tracks, each additional track was introduced with an opt-in question to allow participants to complete only the sections of the questionnaire that they chose to take part in. The survey was distributed electronically and completed online. All postings or emailed survey links were

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

accompanied by an informational invitation to participate, detailing the approximate time commitment involved in participating, and the general topic of the survey, as well as contact information for both the primary investigator and the author. The invitation provided a link to the questionnaire. Participants were encouraged to email or call the primary investigator or the author, whose contact information was included in the invitation to participate and again in the survey, if they had any questions regarding the survey. Prior to beginning the questionnaire, participants were required to answer a question confirming their consent to participate in the survey. This question was included to ensure that participants considered themselves to be well-informed of all aspects of the study, had no questions for the investigators, and were willing to participate in the study prior to beginning the questionnaire.

Various aspects of the questionnaire, including question design and feature layout were implemented using guidelines and recommendations researched by Dillman, Smyth, & Christian (2014). With only 10 exceptions in the total 65 questionnaire items, items on the questionnaire did not require a participant to respond. Those few questions that did require a response were participation eligibility questions, caseload composition questions (necessary to establish the track of questions presented to participants), and questions related to overall use of standardized tests (necessary to ensure that a participant regularly used standardized tests in clinical practice). In order to prevent participant frustration and fatigue, no progress bar was included in the survey (Dillman et al., 2014). All questions were carefully constructed to ensure that answer choices provided both positive and negative responses and that either non-responses or negative responses were possible for all questions.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

The questionnaire was available to participants from January 25, 2015 to March 15, 2015. The end date of the questionnaire's availability was extended twice due to requests from state associations to enable their members to participate in the study.

### **Questionnaire**

Relevant selections from the questionnaire can be found in **Error! Reference source not found.** The first section of survey questions ensured that only eligible participants took part in the survey. These questions requested that the participant confirm their status as a licensed or certified SLP working primarily in a school setting. Information regarding caseload age/grade distribution was also collected to ensure that later survey questions were relevant to the individual participant's current caseload. If participants' answers to eligibility questions indicated they were not qualified to take part, they were exited from the questionnaire to prevent accidental participation of ineligible persons.

Participant responses to caseload questions in the first block of the questionnaire determined which questionnaire track(s) the participant was asked to complete. Those participants who indicated that the majority of their caseload consisted of elementary students were directed to the elementary track of the questionnaire. Participants who responded that at least 30% of their caseload was composed of elementary students were invited to complete the elementary track of the questionnaire in addition to the track for their primary caseload.

**General Assessment Procedures.** Following the eligibility portion of the questionnaire, participants were asked about their general assessment procedures. Participants were asked to identify the likelihood of using different kinds of non-

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

standardized and standardized assessment procedures, from selections of *very likely*, *likely*, *neither likely nor unlikely*, *unlikely*, *very unlikely*. Assessment procedures discussed included both multi-domain assessment procedures and domain-specific assessment procedures. If participants indicated that they were either *likely* or *very likely* to use non-standardized or standardized domain-specific assessment procedures, two additional questions were asked regarding how these procedures were used in the assessment process.

The portion of the questionnaire examining the selection and use of standardized, norm-referenced tests began with the following case presentation:

“You are preparing to evaluate or re-evaluate an elementary school (1st - 5th grade) child who is suspected of having a language impairment. The child is a native speaker of English with typical hearing and typical nonverbal cognitive skills, has no diagnosed developmental disorders or syndromes (e.g. Down Syndrome, Fragile X Syndrome, Autism Spectrum Disorder, etc.), has no history of brain injury/insult, and oral language is his/her primary mode of communication.”

After this case presentation, participants were asked to answer questions regarding their preferred methods of assessment.

**Standardized Test Preferences.** Next, the questionnaire presented a list 41 formal, standardized, norm-referenced language. Tests listed in the questionnaire primarily targeted children ages six to twelve, although some tests’ age ranges extended beyond this window. A complete list of the tests and their age ranges are included in Appendix B. Tests were only included if published within the last 25 years. When a test

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

had a revisions published within the last 5 years, both the most recent and the previous versions were included. The tests included in the study were primarily drawn from frequently used tests as established by prior studies (Betz, et al., 2013; Caesar & Kohler, 2008; Wilson, Blackmon, Hall, & Elcholtz, 1991), and also included the *Montgomery Assessment of Vocabulary Acquisition* (Montgomery, 2008) and the *Test of Written Language, Fourth Edition* (TOWL; Hammill & Larsen, 2009), which were not included in the previous surveys. Different test versions were combined for some analyses.

The following question accompanied the case presentation and list of tests: “If you received the referral above, which of the following language tests would you be most likely to administer to this child? Please select only 1 test; you will have an opportunity to select additional tests next.” In addition to the tests listed, participants could select the provided “Other” option and enter a test of their own choosing. The second question repeated the list of tests and asked participants to select up to 5 additional tests that they would be likely to use when assessing the child described by the case presentation. This question provided an opportunity for participants to enter up to three additional tests that were not listed. A table of the additional assessments provided by participants is included in Appendix C.

**Factors Impacting Test Selection.** After this, the questionnaire asked participants to rate the impact of various factors on their test selection process. Participants were presented with a Likert scale providing the following ratings: *strongly affected, slightly affected, neither affected nor unaffected, slightly unaffected, strongly unaffected*. The investigators chose the factors studied in this survey through review of the research regarding important qualities of standardized tests as well as conversations with SLPs

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

working in different settings, including professionals currently working in the schools, and both clinical and academic speech-language pathology faculty at Idaho State University. Both test-quality factors and situational factors were included. The factors provided to participants included: (a) the time commitment required to administer (not score) the test, (b) the time commitment required to score (not administer) the test, (c) personal familiarity with the test based on past use, (d) specific recommendations from other SLPs, graduate program faculty, or clinical supervisors, (e) reviews of the test published in professional journals, (f) publisher's description of the test, (g) the cost of the test or test materials, (h) the availability of the test, (i) guidelines in school/facility, (j) guidelines in state and federal laws, (k) psychometric features of the test (such as reliability and validity), and (l) specific measures of diagnostic accuracy (such as mean group differences or sensitivity and specificity) for the test as reported in the examiner's manual or in published research articles. Participants were also asked which of these factors was most and least important in their selection process. They were provided with an option to write in their own factor for both of these questions if none of the listed factors were most or least important to the participant.

One factor that was not included in this list was the ability of a test to target a child's area of deficit. As discussed in the literature review, the skills targeted by a test must be carefully evaluated. This is particularly true when assessing students with possible SLI, as children with SLI present with a wide variety of language deficits. However, the case presentation given in this questionnaire was deliberately designed to provide a general description of a student with SLI. This allowed participants' responses to reflect their typical test selection process. Asking participants if their test selection

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

process was affected or unaffected by the skills targeted by the test would result in responses that cannot be effectively analyzed because the degree to which a test may target deficits varies both by test and by presentation of reported language difficulties.

The final section of the questionnaire collected general demographic information, including the region in which the SLPs worked, what degree(s) they held, their certification and/or licensure, and number of years they have been practicing.

### **Participants**

Investigators recruited participants for the survey by contacting ASHA-affiliated state organizations. Organizations were asked to distribute the survey to their membership via email or to post a link to the survey on their website. All 50 state associations were contacted via email and 24 state associations agreed to participate in distributing the survey. Of the remaining states, 13 responded that the request had been passed on to organization leadership, but never indicated a final decision regarding distribution, 8 did not respond to repeated communication attempts, and 4 declined to participate. In addition to distribution through state associations, investigators posted an invitation to participate and a link to the survey on the ASHA Facebook page and on the *ASHA Communities* websites for special interest groups (SIGs), SIG 1: Language Learning and Education and SIG 16: School-Based Issues.

In order to be eligible to participate in the study, SLPs had to be working as speech-language pathologists with a workload that consisted of at least 41% school-based clients. SLP aides and/or assistants were not eligible to participate. A total of 680 participants opened the online questionnaire and 593 participants submitted a questionnaire, resulting in an 87% submission rate. There were 73 participants who

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

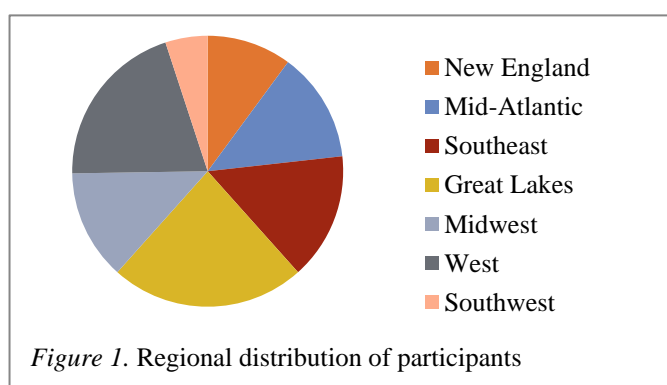
submitted questionnaires but were ineligible to take part in the survey for reasons including refusal to provide consent to participate (2), indication of a profession other than SLP (16), or disclosure of insufficient caseload in a school-based setting (55). These participants were exited from the questionnaire, resulting in a submitted but not completed questionnaire. Therefore, out of the total number of submitted questionnaires, 520 participants submitted completed questionnaires, resulting in a 76% completion rate. Of the participants who completed the elementary track of the questionnaire, 194 completed only the elementary section, 171 completed two different grade-level tracks, and five completed all three tracks. Of the SLPs who completed the elementary track, 332 had caseloads that consisted primarily of elementary school (1<sup>st</sup> through 5<sup>th</sup> grade) students. The remaining participants had caseloads that consisted primarily of preschool and kindergarten students (30 participants) and middle and/or high school (6<sup>th</sup> through 12<sup>th</sup> grade) students (eight participants), although at least 30% of their caseload consisted of elementary students. There were 370 participants who completed the elementary grades portion of the questionnaire.

**Licensure.** The majority of participants, 359 total, were licensed or certified speech-language pathologist, but 10 participants were working with temporary or conditional licensure or certification and one participant was working without licensure or certification. Of the total elementary section participants, 96% of participants held a Certificate of Clinical Competence from the American Speech-Language-Hearing Association, 80% had a state-issued educational provider's license, 61% had a state-issued health provider's license, and 2% had a state-issued working waiver and/or temporary/conditional license. Only one participant reported working with a bachelor's

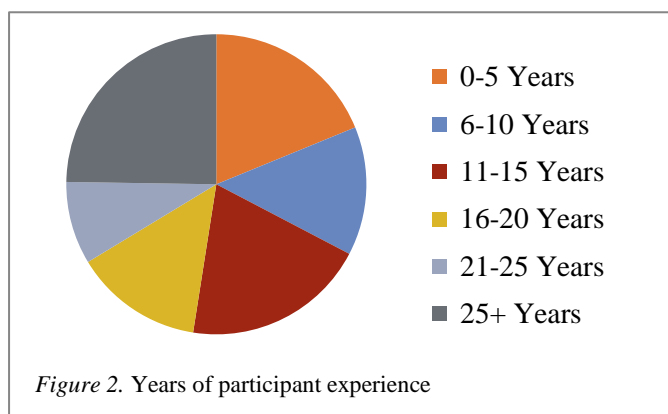
## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

degree, while 96% of participants reported having a master's degree and 4% reported having a doctoral degree.

**Region.** Of the 370 elementary section participants, 367 study participants reported the region in which they currently work. Figure 1 illustrates the proportion of responses by region. The Great Lakes region had the highest response rate of 23%, while the Southwest had the smallest percentage of participants at 5%. See Appendix D for a more detailed listing of states by region.



**Years of Experience.** Number of years of experience ranged from 0 to 25 or more years. Figure 2 illustrates the percentage of responses in each range. The group with the highest response rate had 25 or more years of experience, with a quarter of the responses falling into this range. The group with the lowest response rate had between 21 to 25 years of experience with 9% of the responses.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### Results

Analyses of data from this survey included both descriptive and relational analyses. Basic percentages were calculated to describe SLPs' current assessment procedures.  $\text{CHI}^2$ , cluster analyses, correlation, and multivariate analysis of variance (MANOVA) procedures were used to explore relationships between the variables. The analyses conducted and their results are discussed below.

#### General Assessment Procedures

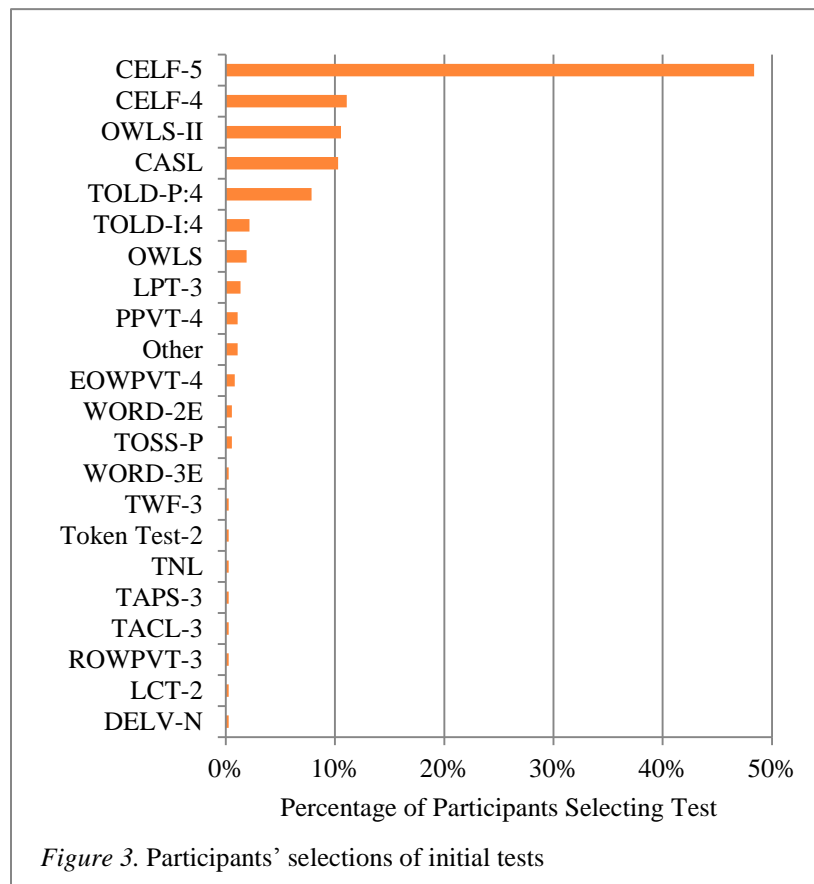
Participants were asked to indicate their likelihood of using different assessment procedures during their evaluation process (see Appendix A, questions 8-13). Responses to these items indicated that a large majority of participants would be *likely* or *very likely* to use standardized norm-referenced multi-domain tests as a part of their evaluation process, with 90% of participants indicating that they would be *very likely* to do so and 8% of participants responding that they would be *likely* to do so. Regarding the selection of domain-specific tests (e.g., tests targeting semantics, morphology, syntax, or pragmatics), 84% of participants indicated that they would be either *likely* or *very likely* to use such a test as a part of their evaluation process. Participants indicated that when they used domain specific assessment procedures, 95% administered these types of procedures in addition to multi-domain-type procedures. Of those that used domain-specific assessment procedures with multi-domain assessment procedures, either standardized or non-standardized, 66% reported using these procedures to increase the information they had about areas of language already assessed, while 34% reported using these assessments to gain information about areas that were not yet assessed.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### Initial Test Selection

Participants were asked to select one (and only one) test for initial assessment of a child given the case presentation included in the questionnaire (see Appendix A, question 14). Percentages for most frequently selected initial tests were calculated and then used to determine the domains most frequently assessed by initial tests.

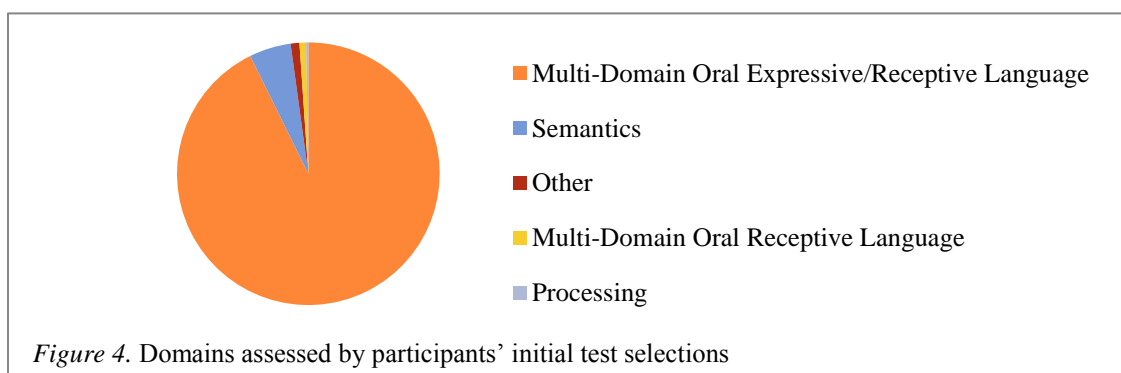


Respondents to this survey most frequently selected multi-domain language tests as their assessment tool of choice. Of the top 10 most frequently selected tests, seven are multi-domain language tests. The *Clinical Evaluation of Language Fundamentals 5* (CELF-5; (Semel, Wiig, & Secord, 2013) was selected by 48% of participants, by far the most frequently selected initial standardized test among the provided options.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Of 41 different tests listed, only 22 tests were selected by participants for primary or initial assessment, as illustrated in Figure 3. As previously mentioned, participants were given an opportunity to select “Other” and enter their preferred test in the provided blank. One “Other” selection indicated that either the *Test of Language Development – Primary: 4* (TOLD-P:4; Newcomer & Hammill, 2008) or the *Test of Language Development – Intermediate: 4* (TOLD-I:4; Hammill & Newcomer, 2008) would be administered, depending on the age of the participant, even though these tests were in the list. Two “Other” selections did not include a test recommendation and the remaining “Other” response indicated more information on the child’s deficits was needed to select a test. A complete list of the tests provided in the questionnaire’s elementary section with abbreviations is included in Appendix B.

As can be seen in Figure 4, the domains frequently assessed by initial tests primarily evaluate multiple domains of oral language both expressively and receptively. Few tests that targeted only a single domain of language were selected, but those single-domain tests that were selected focused on semantics.



A  $\text{CHI}^2$  analysis was used to examine the differences in initial test selection by region and years of experience. The seven most frequently selected tests were identified

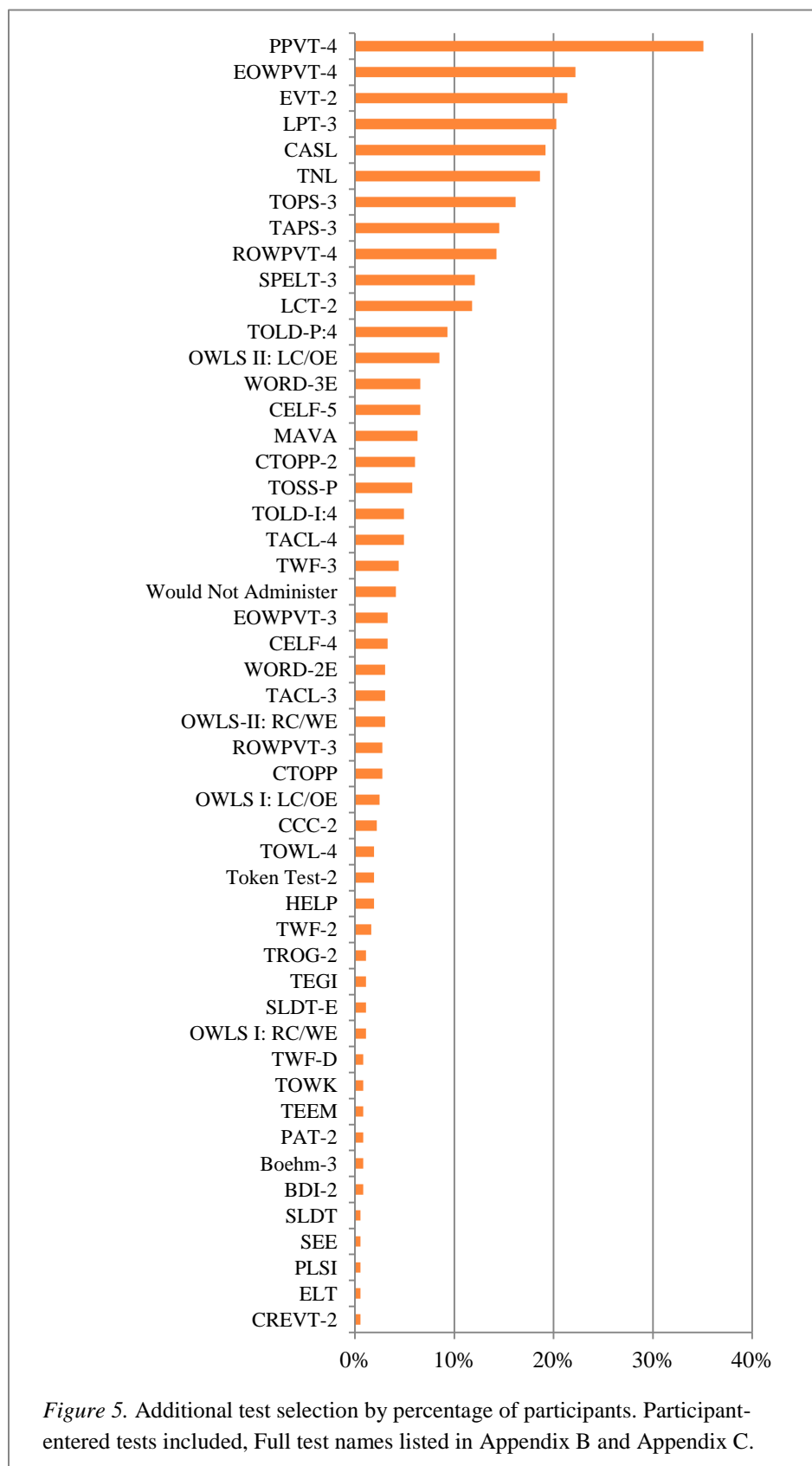
## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

and test editions were combined. After this, tests selected by less than 35 respondents were eliminated to allow for a valid  $\chi^2$  analysis. Initial test selection did not differ significantly for either region ( $\chi^2 (12, N = 302) = 17.95, p = .117$ ) or years of experience ( $\chi^2 (10, N = 304) = 6.34, p = .786$ ). However, although the  $\chi^2$  analysis of the top three (editions combined) initial tests selected by region did not reach a p-value less than .05, the p-value of .117 suggests a possible regional difference. Closer inspection of the data revealed that the *CASL* (Carrow-Woolfolk, 1999) was not selected by any clinician in the New England region, while it was selected at least some of the time in every other region.

### **Additional Test Selection**

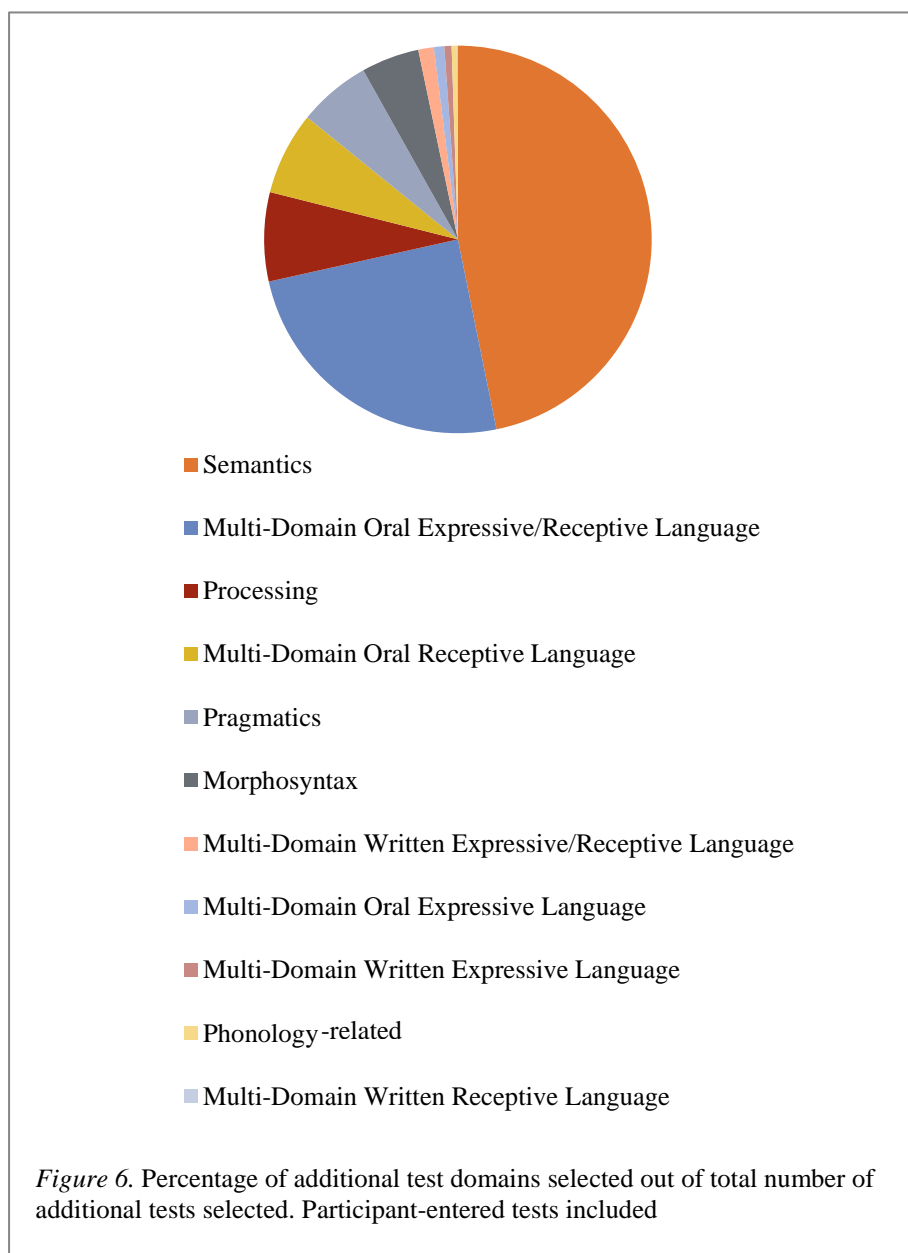
Additional tests selected by participants were analyzed in a similar way to initial tests, as additional tests selected and the domains assessed by these tests were described using percentages. Fourteen participants indicated that they would not administer additional tests and 40 participants elected to write in additional standardized tests they would administer. Additional tests recommended by participants can be found in Appendix C.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Of the types of tests most frequently selected to administer in addition to initial standardized tests, single-word vocabulary measures made up half of the top ten selections, as can be seen in Figure 5. The remaining tests in the top ten selections include multi-domain language test and several tests targeting domain- domain-specific skills, such as morphosyntax, and phonological processing, as well as an evaluation of narrative skills.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

As vocabulary measures, primarily single-word vocabulary tests, were selected by a large percentage of participants to administer as additional tests, it is no surprise that the language domain focus of these additional tests was dominated by semantics, followed by multi-domain oral expressive/receptive language, and then processing tests, as can be seen in Figure 6. Tests examining pragmatics and morphosyntax follow these in much smaller percentages. Multiple related-samples nonparametric analyses using Cochran's Q revealed significant differences between the selection of domain specific tests for Semantics and Morphosyntax ( $Q(1)=209.61, p=.000$ ), Semantics and Pragmatics ( $Q(1)=203.36, p=.000$ ), and Semantics and Processing ( $Q(1)=176.42, p=.000$ ).

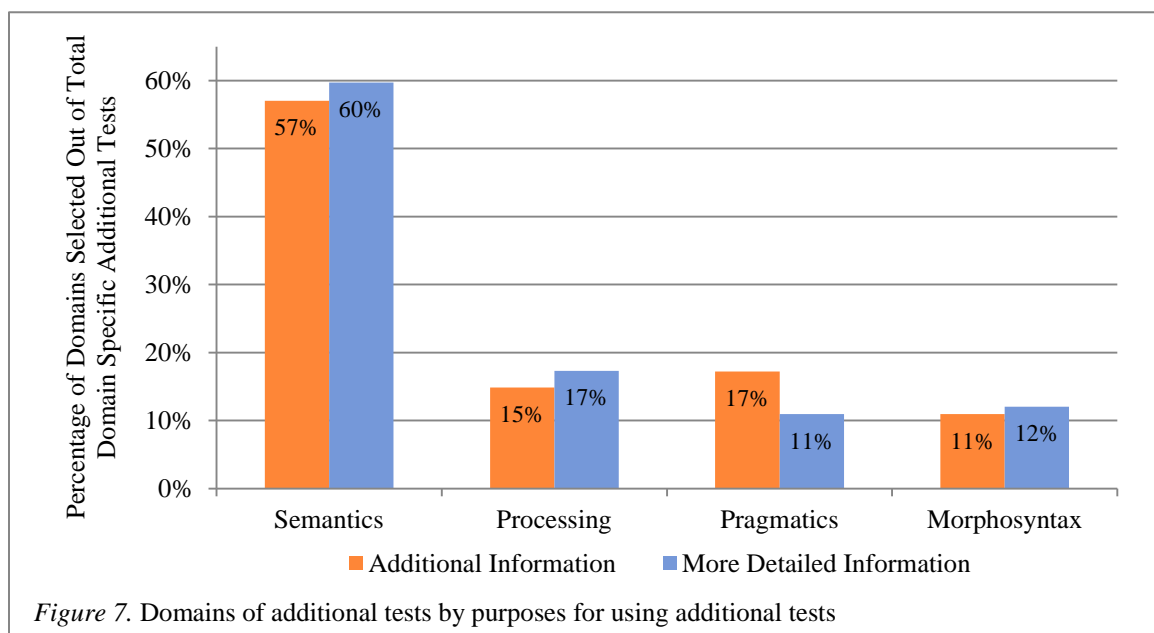
Further analysis was conducted on the additional test selection responses, with participants grouped by their response to Question 13:

“When you use an assessment tool that evaluates a specific area of language in addition to administering one that evaluates many areas of language, you are most likely to administer it... (a) In order to add information about different areas not addressed by the initial assessment, (b) In order to add more detailed information about areas already addressed by the initial assessment”

The domains of the additional tests selected by participants who chose answer (a) in the above question, (that is they use domain specific assessment procedures to add information about areas not addressed by the initial assessment) were compared to the domains of the additional tests selected by participants who chose answer (b). As can be seen in Figure 7, tests targeting semantics are most

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

commonly selected regardless of the participant's stated purpose for using additional tests.



Because of the high percentage of semantics tests selected, these tests were categorized according to type of response required for further examination. These tests were divided into two categories: those requiring only a single word or pointing response, essentially a labeling or identification task, and those requiring a response consisting of a description, definition, or identification of a relationship, which would test deeper semantic knowledge. Those participants who use additional tests to gain new information selected 71% labelling-type semantics tests and 29% deeper-knowledge-type semantics tests. These percentages differed very little from those SLPs who use additional tests to gain more detailed information in areas already assessed, who selected 65% labelling-type semantics tests and 35% deeper-knowledge-type semantics tests.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### Factors Impacting Test Selection

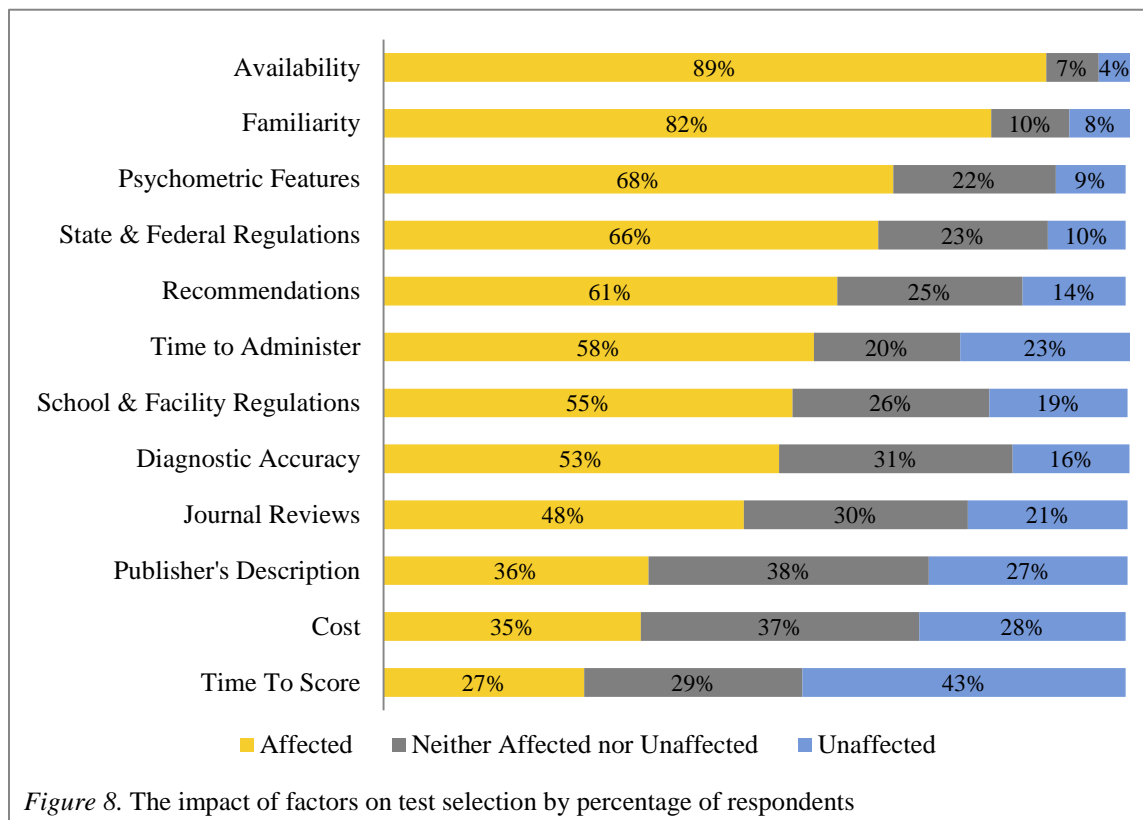
Participants were asked rate the impact of 12 factors on their test selection process using a five-point Likert scale (*strongly affected, slightly affected, neither affected nor unaffected, slightly unaffected, strongly unaffected*). This five-point scale was then condensed to a three-point scale (*affected, neither affected nor unaffected, unaffected*), described by percentage of participants' rating selections, and analyzed using several different tests.

As can be seen in Figure 8, the greatest number of participants indicated that factors of test availability, personal familiarity, and psychometric features were of greatest importance when selecting a test to administer. Time to score the test was selected as the least affecting factor by the greatest number of participants, closely followed by test cost and the publisher's description of the test.

To evaluate possible relationships between the factors, several different analyses were conducted. Analyses of factor relationships included a hierarchical cluster analysis to explore participant patterns of test selection and a nonparametric correlation using Spearman's  $\rho$  to evaluate the relationships between the factors themselves. Additionally, a MANOVA was used to evaluate the influence of test selection factors and their potential relationship to geographic region and years of experience.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS



The factors were analyzed with a hierarchical cluster analysis using Euclidian centroid method to explore if participants selected factor importance in distinct patterns. A single-cluster solution was found, indicating that particular factors did not cluster together in subgroups and therefore no patterns of factor ratings were found among participants.

A nonparametric correlation using Spearman's  $\rho$  was used to examine the relationships between the test selection factors. In Ferguson (2009), a primer in evaluating effect sizes in social sciences,  $\rho$  is considered to indicate a minimum practically significant effect size at  $\rho = .200$  and  $\rho = .500$  describes a moderate effect size. As many of the factors in this study demonstrated practically significant effect sizes,

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Table 1 lists only those factors that correlated with a  $\rho$  in the range of a moderate effect ( $\rho > .400$ ).

Table 1

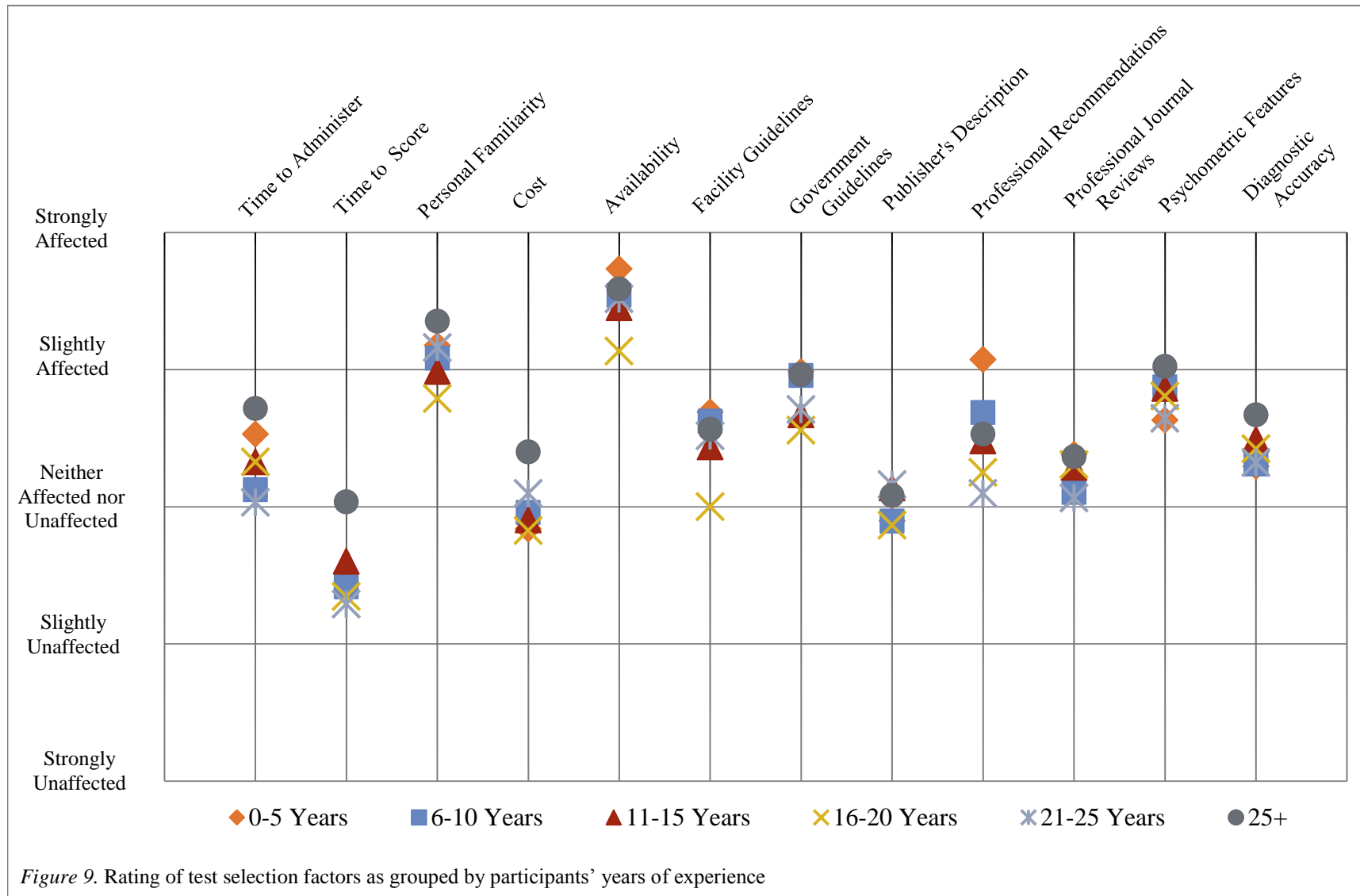
*Significant Spearman's  $\rho$  Correlations of Test Selection Factors*

	$\rho$
Diagnostic Accuracy & Psychometric Features	.657*
Facility Guidelines & Government Guidelines	.598*
Time to Score & Time to Administer	.596*
Journal Reviews & Psychometric Features	.464*
Journal Reviews & Diagnostic Accuracy	.425*

\* $p < .001$

MANOVA was conducted to investigate any possible differences in test selection factor ratings by region, but results from this test were non-significant. A MANOVA was also used investigate any differences in test selection factor ratings by years of experience cohorts. Because of unequal variance between the groups (as indicated by a significant Box's test), Pillai's Trace was used to evaluate the significance of the MANOVA (Tabachnick & Fidell, 2012). Pillai's trace indicated significant differences between the groups ( $p < .001$ ). Subsequent between-subjects ANOVAs indicated that differences occurred between cohorts for factors *time to score* ( $p = .009$ ), *personal familiarity* ( $p = .037$ ), *availability* ( $p = .013$ ), *facility guidelines* ( $p = .048$ ), and *professional recommendations* ( $p < .001$ ). These differences can be visualized in Figure 9.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Post-hoc testing using a least significant difference (LSD) test was conducted to identify where significant differences occurred between groups. The significant differences are listed in Table 2. Notable differences were found between SLPs with 25+ years of experience and all other cohorts on the test selection factor *time to score*, as this factor was rated more highly by the more experienced cohort than by any other group. Another interesting difference was found for professionals with 0-5 years of experience, who indicated the factor *recommendations from other professionals* has more influence on their test selection process than any other group.

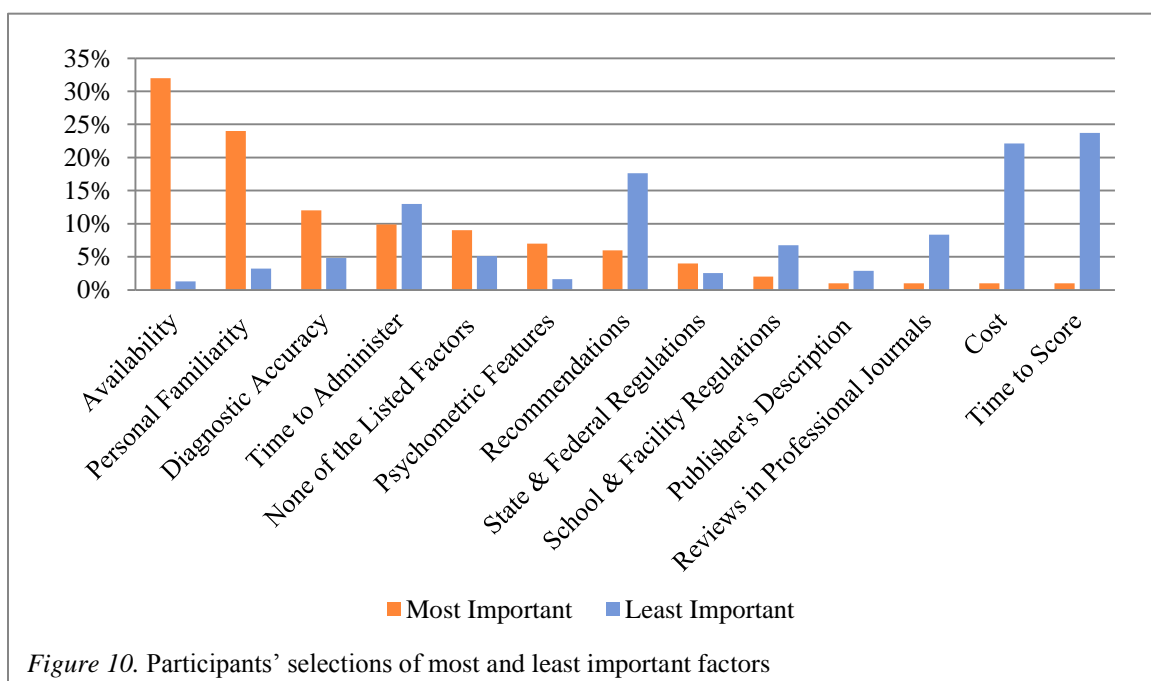
Table 2  
Multiple Comparisons: Years of Experience and Test Selection Factor Ratings

Factor	(I) Years of Experience	(J) Years of Experience	<i>p</i>
Time to Score	25+ Years	0-5 Years	.005
		6-10 Years	.008
		11-15 Years	.035
		16-20 Years	.002
		21-25 Years	.006
Personal Familiarity	16-20 Years	0-5 Years	.035
		25+ Years	.001
	25+ Years	11-15 Years	.022
Availability	16-20 Years	0-5 Years	<.001
		6-10 Years	.022
		11-15 Years	.049
		25+ Years	.004
Facility Guidelines	16-20 Years	0-5 Years	.002
		6-10 Years	.011
		11-15 Years	.048
		25+ Years	.009
Recommendations	0-5 Years	6-10 Years	.048
		11-15 Years	.001
		16-20 Years	<.001
		21-25 Years	<.001
		25+ Years	.001
	6-10 Years	16-20 Years	.035
		21-25 Years	.014

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

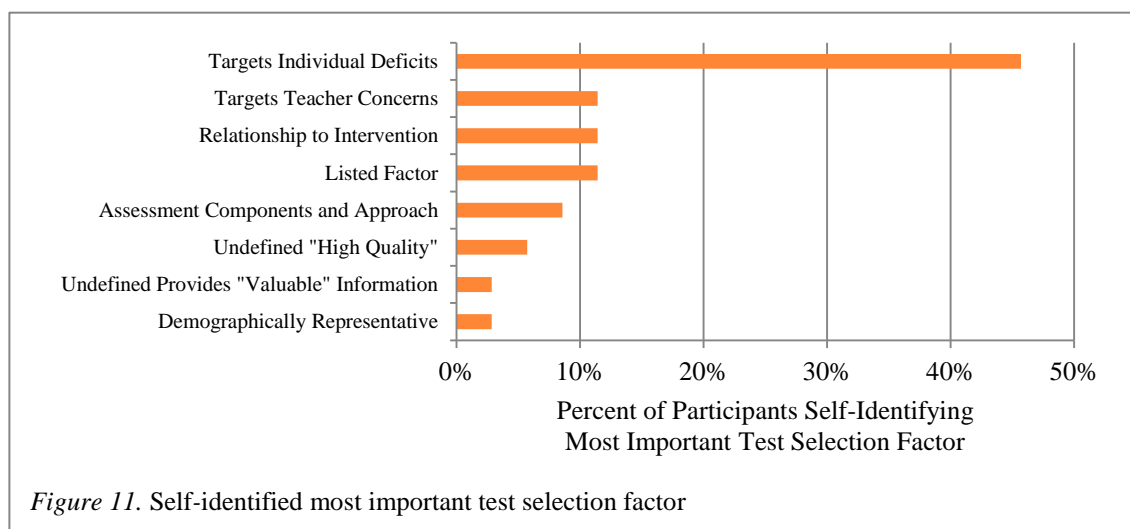
### Most and Least Important Factors

After they were asked to rate test selection factors using the scale, participants were also asked to indicate which factors were most and least important to them during their assessment process (Appendix A, questions 17-20). These responses were described using percentages and a  $\text{CHI}^2$  test was calculated to examine the relationship between selected factors and initial test selection. The responses, shown in Figure 10 indicate that the factors of *availability* and *personal familiarity* were selected by a majority of participants as most important, and the factors of *administration time* and *diagnostic accuracy* were identified more often than *psychometric features*. Of the least important factors, *scoring time* and *cost* remained the top two least important factors.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

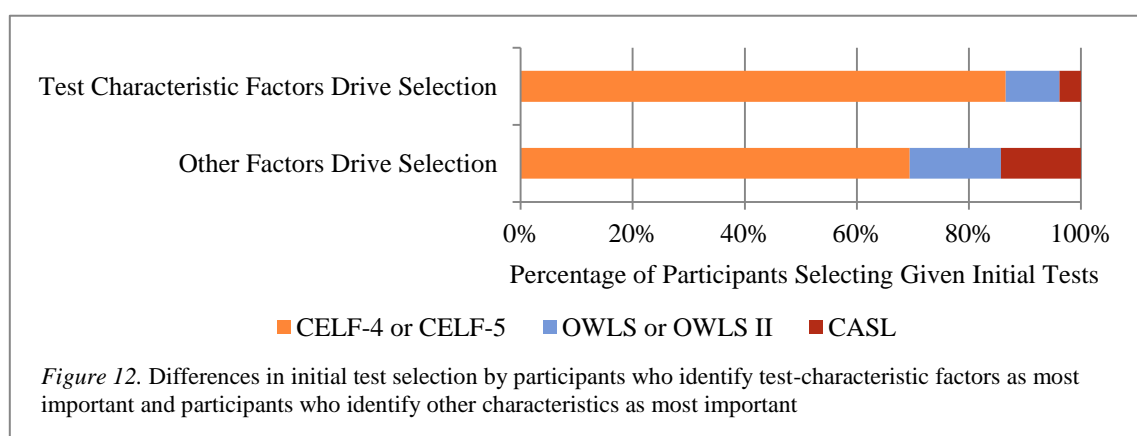
Participants were given the option of naming their own most or least important factor. The self-identified most important factors, 28 in total, are listed in Appendix E. These factors were categorized and evaluated using percentages as shown in Figure 11. More than half of the participants that identified a most important factor indicated that the test's ability to target a client's areas of deficit was the feature that had the greatest impact on whether or not the participant would select a test for administration. Only eight participants self-identified least important test selection factors. Of the eight, four wrote in previously listed factors, two were unsure, one felt that all factors were important, and one felt that the value of the information provided by the initial test would be a determining factor in the selection process.



Investigators grouped responses from participants based on whether they identified *test-characteristic factors*, such as reliability, validity, and diagnostic accuracy, as most important or *other factors* as most important. The initial test selection of the *test-characteristic* group was compared with the initial test selection of participants who identified *other factors* as most important, as can be seen in Figure 12. A  $\text{CHI}^2$  test

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

demonstrated that participants who selected *test-characteristic* factors as their most important factor were more likely to select a different initial test than the other participants ( $\chi^2 (2, N = 304) = 6.76, p = .034$ ). Those who identified *test-characteristic* factors as most important were more likely to select the fourth or fifth edition of the *CELF* (Semel, Wiig, & Secord, 2003; Semel, Wiig, & Secord, 2013) than either edition of the *OWLS* (Carrow-Woolfolk, 1996; Carrow-Woolfolk, 2011) or the *CASL* (Carrow-Woolfolk, E., 1999).

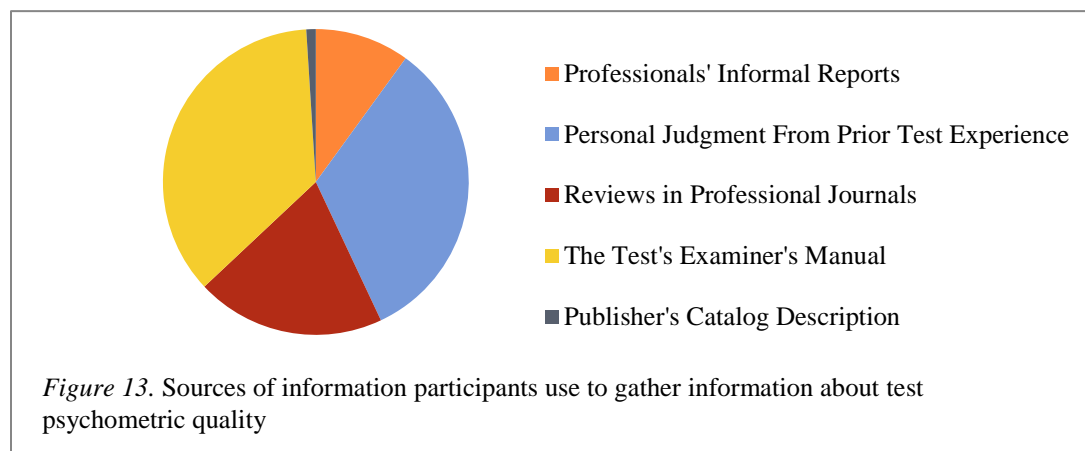


### Sources of Information Regarding Test Psychometric Quality

Finally, the resource from which participants gain information about tests' psychometric qualities was described using percentages. Many participants reported that, when determining the psychometric quality (such as the test's reliability and validity) of a test, they look to the test's examiner's manual for information. However, as can be seen in Figure 13, almost as many participants rely on their personal experience with a test to make determinations about the psychometric qualities of a test. As the number of participants indicating that the publishers' catalog descriptions were a primary resource was extremely low, the following analyses were conducted with the participants divided into groups based on their primary source of psychometric information. These were

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

professionals' informal reports, personal experience, reviews in professional journals, and examiner manuals.



The differences between groups on initial test selection and test selection factor ratings were analyzed in several ways, including a  $\chi^2$  and a MANOVA. A  $\chi^2$  was used to examine the relationship between sources of information and initial test selection, but no significant differences were found ( $\chi^2(6, N = 297) = 7.71, p = .260$ ).

Investigators grouped participants based on their identified primary source of test psychometric information. Using these groups, a MANOVA was conducted investigate any differences in test selection factor ratings between the groups. Because of unequal variance between the groups, Pillai's Trace was used to evaluate the significance of the MANOVA (Tabachnick & Fidell, 2012). Pillai's trace indicated significant differences between the groups ( $p < .001$ ). The results of the MANOVA can be seen in Figure 14.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

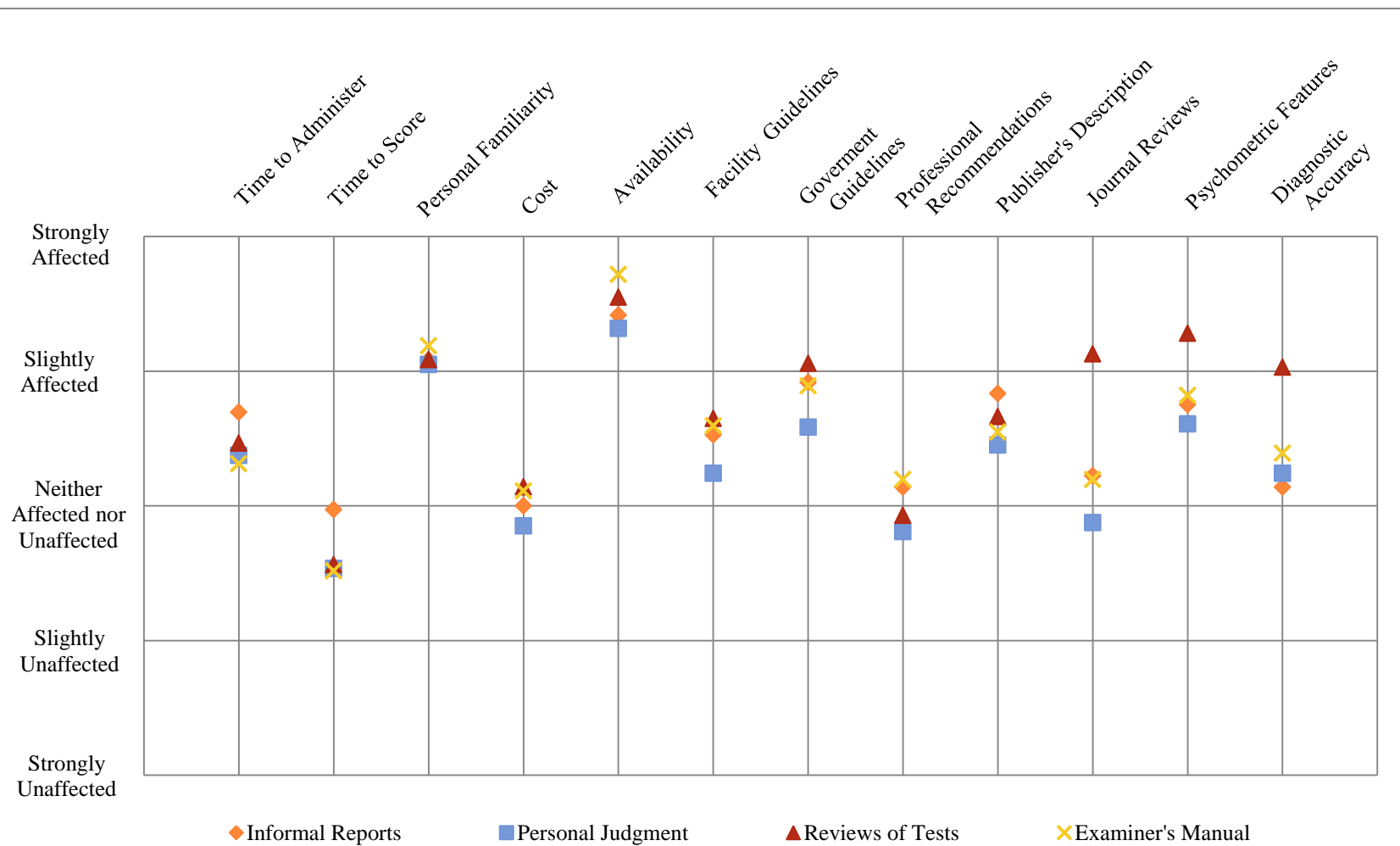


Figure 14. Rating of test selection factors as grouped by participants' primary sources of psychometric information

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Subsequent between-subjects ANOVAs indicated that differences occurred between cohorts for factors *availability* ( $p = .004$ ), *government guidelines* ( $p = .026$ ), *publisher's description* ( $p = .032$ ), *journal reviews* ( $p < .001$ ), *psychometric features* ( $p < .001$ ), and *diagnostic accuracy* ( $p < .001$ ). These differences between groups can be seen in Figure 14.

Post-hoc testing using a LSD test was conducted to identify where significant differences occurred between groups. Significant differences are listed in Table 3. Of the significant differences between the groups, the most notable difference occurs for participants who indicated that they rely on journal reviews of tests for psychometric information. These participants rated *journal reviews*, *psychometric features*, and *diagnostic accuracy* as factors that more strongly impact their test selection decision than did participants in other groups.

Table 3

Multiple Comparisons: Source of Psychometric Information & Test Selection Factor Ratings			
Test Selection Factor	Source of Psychometric Information	Source of Psychometric Information	<i>p</i>
Availability	Personal Judgment	Examiner's Manual	<.001
Federal Guidelines	Personal Judgment	Reviews of Tests	.005
		Examiner's Manual	.031
Publisher's Description	Personal Judgment	Examiner's Manual	.005
Journal Reviews	Reviews of Tests	Informal Reports	<.001
		Personal Judgment	<.001
		Examiner's Manual	<.001

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Psychometric Features	Reviews of Tests	Informal Reports	.010
		Personal Judgment	<.001
		Examiner's Manual	.002
Diagnostic Accuracy	Reviews of Tests	Informal Reports	<.001
		Personal Judgment	<.001
		Examiner's Manual	<.001

---

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### **Discussion**

As SLPs are the primary professionals responsible for diagnosing and treating language-based disorders such as SLI, thoughtful decision-making during the assessment process is critical. The goal of this survey was to explore the factors currently driving these decision-making processes. By providing a greater understanding of current practices, this study allows professionals to identify possible areas for future changes to improve outcomes, both for SLPs in the schools and their students.

### **Test Selection**

Past research has indicated that a majority of SLPs do use standardized tests as a part of their assessment process. The results from this survey, indicating that 98% of participants would be *very likely* or *likely* to make use of a standardized test in their evaluation process, underscores the value of recognizing how and why these tests are selected. Findings from the test selection portion of the survey also correspond with findings from past surveys (Betz et al., 2013; Caesar & Kohler, 2009; Wilson et al., 1991), indicating that the most frequently administered tests are multi-domain language tests and semantics tests.

As the current survey asked not only what tests would the participants be likely to give, but which tests they would give first and which they would use in addition to the first test, more detailed information is now available regarding the sequence used by school-based SLPs who participated in the survey. Responses indicated that clinicians most frequently select multi-domain tests as their initial evaluation choice and often use additional tests of semantics or, to a lesser extent, other domain specific tests to supplement the results from the multi-domain test. Administering tests in this sequence

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

could enable SLPs to gain a broader view of a child's language abilities and then more narrowly focus subsequent test selections. Responses to other items on the questionnaire, however, might suggest that this administration sequence may not accurately reflect the intentions of the SLP conducting the assessment.

In the general assessment procedure section, SLPs were asked to indicate why they administered additional domain-specific tests, either to add information in a language domain they have already assessed or to gain information about a language domain they have not yet assessed. However, the domains evaluated by the additional domain-specific tests that participants selected did not appear to be different, regardless of the SLPs' stated purpose. Unfortunately, it was not possible further examine this discrepancy between test purpose and test selection. This was because the survey did not ask participants which subtests of any selected initial multi-domain tests they would likely administer. Therefore, it was not possible determine whether the additional tests that participants selected actually assessed different domains than the test selected for initial administration. It is possible that the lack of difference in test selection between the groups may be a result of the generic case presentation described in the questionnaire. However, the lack of difference between groups is suggestive of a degree of incongruence between SLPs' intended purpose for follow-up assessment and the actual information gained by this assessment.

Another surprising outcome of the survey results was the small number of tests selected for initial assessment. With 41 different tests to choose from, a large majority selected one of only seven different tests. The high percentage of SLPs who selected the *CELF-5*, arguably one of the psychometrically stronger tests available, as their preferred

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

initial test may indicate that SLPs are attending to the reliability, validity, and diagnostic accuracy of standardized tests. However, results from the survey indicate that 32% of clinicians rate availability as the most important factor in test selection, as compared to 7% of SLPs who rate psychometric features as most important, and 12% of SLPs who rate diagnostic accuracy as most important.

It is notable that the psychometric features of the top seven most frequently selected initial tests are not the same, suggesting that psychometric features may not be the most important test-selection factor for many SLPs. When all of these top tests (the *CELF-5*, the *CELF-4*, the *Oral and Written Language Scales, Second Edition [OWLS-II]*; Carrow-Woolfolk, 2011], the *CASL*, the *TOLD-P:4*, the *TOLD-I:4*, and the *Oral and Written Language Scales [OWLS]*; Carrow-Woolfolk, 1996]) are scrutinized, the examiner's manuals of both the fourth and fifth editions of the *CELF* and of the *TOLD I:4* report adequate sensitivity and specificity levels, that is, sensitivity and specificity levels of 0.8 or greater (Plante & Vance, 1994). The sensitivity and specificity levels of the other tests are either inadequate or not addressed at all (Betz et al., 2013). When this variation of quality is considered in concert with how highly *availability* was rated by participants, it is possible that the *CELF-5* has simply become more widely available to school-based SLPs than the other tests, and therefore is used more frequently. Alternately, it is possible that it is the most available test because it is deemed to have good psychometric characteristics, although that was not reflected in the study participants' responses.

That being said, SLPs who rated test-quality factors (such as *reliability*, *validity*, and *diagnostic accuracy*) as having a greater impact on their test selection process were

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

more likely to choose a version of the *CELF* than other tests. Therefore the prioritization of factors does appear to influence test selection for some clinicians.

When examining the selection of tests chosen in addition to an initial test, it is clear that semantic tests, in particular single-word vocabulary tests, are the most commonly selected additional tests. While single-word identification and naming tasks can provide some information regarding vocabulary recognition and/or naming ability, they provide less information about semantic network organization and word-learning abilities. Although it was expected that one-word vocabulary tests would be frequently chosen as additional tests to administer, the overwhelming selection of these tests was unexpected, as the case presentation provided in the questionnaire described, in very general terms, a student with possible SLI. While many children with SLI demonstrate difficulties with semantic skills, the nature of this deficit is complex, relating to multiple aspects of word learning, storage, and recall (see Leonard, 2014, for a review) and not just single word naming or identification. Brackenbury and Pye (2005) recommend more dynamic word-learning or word-relationship tasks as being more effective than one-word/pointing response semantics tests in identifying these deficits in children with SLI. Another factor that may have influenced this choice is that the case description did not specifically indicate SLI. Thus, many clinicians may consider basic recognition and naming vocabulary to be key factor in their diagnostic process.

Domain specific tests targeting morphology and syntax were selected infrequently, which was somewhat surprising. Deficits in morphology (particularly tense and agreement morphemes) and syntax are hallmark deficits for children with SLI (Bedore & Leonard, 1998; Charest & Leonard, 2004; Conti-Ramsden, 2003; Leonard,

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

1995; Leonard, 2014a; M. A. Nippold, Mansfield, Billow, & Tomblin, 2008; Rice & Wexler, 1996; Rice, Wexler, Marquis, & Hershberger, 2000; Schuele & Dykes, 2005; Scott & Stokes, 1995). Because the relationship between SLI and morphological and syntactic deficits is well-known, it was expected that more of these tests would have been selected, given the case presentation in the questionnaire. However, this is understandable considering that there are few existing tests that primarily target morphology and syntax. Of the 41 tests presented in the questionnaire, only the *Structured Photographic Expressive Language Test, Third Edition* (SPELT-3; Dawson, Stout, & Eyer, 2003), the *Test of Early Grammatical Impairment* (TEGI; Rice & Wexler, 2001), and the *Test for Reception of Grammar, Second Edition* (TROG-2; Bishop, 2003) are designed to target these domains. Additionally, there are no domain specific tests that focus on comprehension or production of complex syntax. While these domains may be targeted in subtests on multi-domain tests, most subtests have too few exemplars of grammatical morphemes or specific complex syntactic structures to enable an SLP to make any firm conclusions about the nature of a child's morphological and syntactic skills.

Another possible factor potentially related to the limited use of tests targeting morphosyntax is that SLPs themselves may not feel confident in their morphosyntactic knowledge and therefore may inadvertently overlook more rigorous assessment of these domains. In a survey of first-year SLP students in 10 different graduate programs, Blackley and McCready (2006) found that students' mean number of correct responses for morphology tasks (including counting morphemes, and identifying free, bound, inflectional, derivational, and grammatical morphemes) was 21 out of 31 items (68% accuracy). Students' mean accuracy on syntactic items (including identifying sentence



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

types and parts of speech) was 5.3 out of 14 items (38% accuracy). Although the participants in this study were first-year SLP graduate students, their markedly low performance on basic morphological and syntactic tasks indicates serious area of weakness that may or may not be sufficiently remediated during graduate studies. If SLPs are not confident in their knowledge of morphology and syntax, they may unintentionally avoid assessing these areas, a possibly damaging oversight for children with SLI.

The results from the test selection portion of the survey provide important information regarding which standardized language tests are most often selected by school-based SLPs, but this study endeavors to move beyond the ‘what tests are used’ aspect of assessment. The following exploration of the responses to the test selection factors portion seeks to provide more information regarding why SLPs select the tests they do.

### **Factors Impacting Test Selection**

As the selection of standardized language tests is a complex and multi-faceted process, it was not surprising that many of the factors impacting this process correlated with each other. However, not all correlations were notable. Positive correlations between some situational factors, such as *time to score* and *time to administer*, have obvious connections, such as a high value placed on time. The correlation between the test-quality factors of *diagnostic accuracy* and *psychometric features* indicates that those who value one aspect of test-quality are likely to value another as well. The factor *journal reviews*, which correlates positively with the expected factors of *psychometric features* and *diagnostic accuracy*, also correlates with *publisher’s description*. This may suggest a person who regards the publisher’s description as an important information source is also

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

likely to seek more information on the test using published reviews. The situational factor correlation between *availability* and *facility guidelines* suggests that tests perceived as providing information necessary to comply with facility guidelines may be more readily available to SLPs.

While the rating of most and least important factors did not contradict the Likert scale ratings, the additional factors added by participants indicate that selecting a test that appropriately targets a child's reported area of deficit is an important consideration for more than half of clinicians that entered their own factor. As was discussed in the literature review, selecting a test that examines perceived areas of deficit is vital to an accurate assessment. This factor could not be included with the rest of the factors discussed in the questionnaire, as it would have been impossible to analyze due to the non-specific nature of the case presentation. Greater detail is provided on the exclusion of this factor in the Methods section (see pp. 25-37). However, its appearance in participants' responses to this question is a heartening sign for evidence-based assessment practices, indicating that clinicians are taking student's individual needs into account when planning assessment.

### **Years of Experience and Factors Influencing Test Selection**

Factors that influence test selection appear to have differing degrees of influence depending on SLPs' years of experience in the field. The relationship between various factors and the years of experience an SLP has provides some interesting information regarding practice priorities. While the situational factor of *personal familiarity* is considered to be slightly important to most cohorts, those SLPs who have worked 25 years or more appear to consider it to be much more important than those who have

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

worked 16 – 20 years. While the situational factor of *availability* is the highest rated factor across all groups, the group for whom *availability* appears to have the lowest impact are clinicians with 16 – 20 years of experience. If these clinicians are, as we noted above, less concerned with personal familiarity with a test, they may also have the seniority which enables them to seek out tests that are not already available to them. Finally, the clinicians who appear to rely most heavily on the situational factor of *professional recommendations* are, predictably, newer professionals with five years of experience or less, while clinicians with greater experience rate recommendations progressively less and less important until they reach 25 years of experience or more. This may suggest that more experienced clinicians feel that their familiar tests might be becoming outdated and are seeking recommendations from their peers for newer tests.

### **Source of Psychometric Information and Factors Influencing Test Selection**

The source an SLP uses to gather psychometric information about a test appears to have some impact on which factors have the most impact on their test selection process. Clinicians who rely on their *personal judgement* for test-quality evaluations rate many factors as being less important. These clinicians may feel that they have a clear understanding of a test's quality, but it is important to keep in mind that expert opinion is a relatively low level of evidence and therefore it is vital to support findings with quality research as well as professional opinion.

In contrast, those participants who reported relying on *reviews of tests* as a source of psychometric information tend to rate both *psychometric quality* and *diagnostic accuracy* as more important factors in their test selection process. However, relying primarily on journal reviews restricts clinicians to a very limited pool of available

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

information. There are very few major reviews that evaluate the psychometric features of available language tests, either for individual tests or across multiple tests. Among the available reviews, those that discuss multiple tests contain limited detail about the tests they evaluate and those that are more detailed provide information on only a few tests. While many smaller reviews are published as poster presentations every year, these are not always easy to find. Therefore, while independent and rigorous reviews of tests are likely to provide better information about the test for consumers, the information available from this resource at this time may be too limited to provide all psychometric information about available tests.

However, as can be seen in the survey results, those who rely on the information in the *examiner's manual* for a test's psychometric information do rate factors differently from both those who rely on *personal judgment* and those who rely on *journal reviews* of tests. These clinicians rate factors like *cost*, *availability*, and *publisher's description* as having a greater effect on their decision-making process than the other groups. This pattern of factors may suggest that this group values efficiency in their practice. However, truly making use of the test examiner's manual to evaluate the many qualities of any test is far from an efficient task. Each publisher may use different manual layouts, different analyses for reliability and validity calculations, and different ways of discussing test features. This makes the process of evaluating tests a challenge and any attempt at fairly comparing tests an arduous and time-consuming task. Altogether, although several different sources of information for test psychometric quality exist, each group's source appears to influence their perspective on many factors, not only psychometric features.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### **Best Practice Recommendations and Factors Impacting Test Selection**

Comparing the results of this survey to preferred practice patterns set forward by ASHA for the characteristics of standardized, norm-referenced tests selected for language assessment, only one of the factors examined in the current research is clearly mandated: “Assessment tools are selected with regard to...evidence of adequate reliability and validity” (p. 47, 2004). In the technical report developed by the National Joint Committee on Learning Disabilities, standardized tests included in the assessment process “must be reliable, valid, and have current normative data” (1994). Given these guidelines, it is interesting to note that only 7% of SLPs selected *psychometric features* (such as validity and reliability) as the most important factor in their decision making process. While best practice recommends that test-quality factors drive the decision-making process when selecting tests, it is clear from the current study that situational factors play a substantially larger role in practice.

Although it would be easy to look at this divergence of preferred practice and actual practice and point accusing fingers, the reality is that clinicians must make these test selection decisions while managing many different professional challenges. Each factor will play a role in the selection decision. How large a role is played, however, will be determined by many different aspects of a SLP’s situation. The challenge brought before us now is how to influence contexts in such a way that those factors that relate to test-quality are, at the very least as salient, if not more salient, than situational factors. By integrating past research with these current findings, it may be possible to identify a few next steps to take to improve how SLPs select tests for children with SLI.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

As was discussed in the literature review, several SLP surveys have indicated that time constraints are a major barrier to implementing evidence-based practice for school SLPs (Hoffman et al., 2013; Meline & Paradiso, 2003; O'Connor & Pettigrew, 2009; Zipoli & Kennedy, 2005). These previous findings are indirectly confirmed by the current study, as the responses gathered from participants indicate that situational factors may often be the driving force behind school-based SLPs' test selection decisions. In a school SLP's busy schedule, the time required to establish the quality of a test can appear daunting or untenable due to the scarcity of professional journal reviews, the labyrinthine design of most examiner's manuals, and the limited information available to SLPs prior to test acquisition. If this information was more readily available, the saliency of test-quality factors would likely increase, resulting in more effective and evidence-based assessment selection.

### **Limitations**

Recruitment of participants for the present survey took place on professional online discussion boards and through professional state organizations. Necessarily, those who had the opportunity to participate in the questionnaire are more likely to be active in their professional community already, leading to a degree of selection bias. Additionally, some factors may be perceived to be inherently more important (such as test reliability, validity, or diagnostic accuracy). As these factors are discussed in preferred practice guidelines and research, participants may have felt that they should rate these factors as having a greater impact on their test selection process than they actually do. Other factors may be perceived as being inherently less important (such as cost of the test or time to score a test), and therefore participants may have rated these factors as less impactful

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

than they actually are. These perceptions of factors' intrinsic value may have led to some self-reporting bias that does not reflect actual practice.

### **Conclusions and Future Directions**

SLPs in schools must take into account both situational and test-quality factors. Time is limited and therefore without easy access to test-quality information, situational factors may be easier to evaluate, resulting in situational factors having a greater impact on the test selection process than the SLP consciously realizes. Although the respondents to the survey indicate that they do attempt to follow some aspects of evidence-based practice by using published reviews of tests or the examiner's manual to establish a test's psychometric information or considering a test's ability to target a student's perceived deficits, the situational pressures of time and limited easy access to test-quality information may limit the practical application of best practices.

Future directions indicated by the current research include identifying clinicians' needs regarding test information access so that test-quality factor information can become as salient as situational factors. It is also clear that more independent reviews of tests are needed to enable more effective selection, both in terms of test-quality and in terms of appropriateness for any particular client. Finally, the few domain specific tests available that directly target morphosyntactic language structure are not often used. As weaknesses in these areas are a hallmark of children with SLI, more education regarding the importance of morphological and syntactic assessment needs to be available to school-based clinicians in order to ensure more efficient and effective assessment of children with SLI.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Although the use of standardized, norm-referenced tests is only a part of a complete assessment process, the selection and use of an appropriate, high-quality test may contribute greatly to the quality of the overall assessment process. Current practice does not always meet best practice standards; however, the quality of standardized tests continues to improve. As professionals work to increase the availability of and access to test-quality information, school-based SLPs can continue to improve their test selection process for children presenting with possible SLI, which in turn can lead to more effective treatment and better outcomes for the children they serve.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Psychiatric Association, & American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). Washington, D.C: American Psychiatric Association.
- American Speech-Language-Hearing Association. (2004). *Preferred practice patterns for the profession of speech-language pathology [Preferred Practice Patterns]*. Rockville, MD: American Speech-Language-Hearing Association. Retrieved from Available from [www.asha.org/policy](http://www.asha.org/policy). doi: 10.1044/policy.PP2004-00191
- American Speech-Language-Hearing Association. (2010). *Roles and responsibilities of speech-language pathologists in schools [Professional Issues Statement]*. American Speech-Language-Hearing Association. Available from [www.asha.org/policy](http://www.asha.org/policy). doi:10.1044/policy.PI2010-00317
- American Speech-Language-Hearing Association. (2012). *2012 Schools survey report: SLP caseload characteristics*. American Speech-Language-Hearing Association. Retrieved from [www.asha.org/research/memberdata/schoolssurvey/](http://www.asha.org/research/memberdata/schoolssurvey/)
- Arndt, K. B., & Schuele, C. M. (2013). Multiclausal utterances aren't just for big kids: A framework for analysis of complex syntax production in spoken language of preschool-and early school-age children. *Topics in Language Disorders*, 33(2), 125–139.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*(5), 1185–1192.
- Betz, S. K., Eickhoff, J. R., & Sullivan, S. F. (2013). Factors influencing the selection of standardized tests for the diagnosis of specific language impairment. *Language, Speech & Hearing Services in Schools, 44*(2), 133–146.
- Bishop, D. (2003). *Test for reception of grammar, second edition* (second). San Antonio, TX: Pearson.
- Bishop, D. V. M. (2010). Which neurodevelopmental disorders get researched and why? *PLoS ONE, 5*(11), 1–9.
- Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems: Terminology for children with language problems. *International Journal of Language & Communication Disorders, 49*(4), 381–415. <http://doi.org/10.1111/1460-6984.12101>
- Bishop, D. V. M., & Clarkson, B. (2003). Written language as a window into residual language deficits: A study of children with persistent and residual speech and language impairments. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 39*(2), 215–237.
- Brackenbury, T., & Pye, C. (2005). Semantic deficits in children with language impairments: Issues for clinical assessment. *Language, Speech & Hearing Services in Schools, 36*(1), 5–16.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- Caesar, L., & Kohler, P. (2009). Tools clinicians use: A survey of language assessment procedures used by school-based speech-language pathologists. *Communication Disorders Quarterly*, 30(4), 226–236.
- Carrow-Woolfolk, E. (1996). *Oral and written language scales*. Torrance, CA: Western Psychological Services.
- Carrow-Woolfolk, E. (2011). *Oral and written language scales, second edition*. Torrance, CA: Western Psychological Services.
- Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language*. Circle Pines, MN: AGS.
- Center for Disease Control and Prevention. (2008). Prevalence of Autism spectrum disorders — Autism and developmental disabilities monitoring network, 14 sites, United States, 2008. *Morbidity and Mortality Weekly Report Surveillance Summaries*, 61(3). Retrieved from <http://www.cdc.gov/mmwr/pdf/ss/ss6103.pdf>
- Charest, M. J., & Leonard, L. B. (2004). Predicting tense: Finite verb morphology and subject pronouns in the speech of typically-developing children and children with specific language impairment. *Journal of Child Language*, 31(1), 231–246.
- Cohen, J. (1988). Statistical power for the social sciences. *Hillsdale, NJ: Laurence Erlbaum and Associates*.
- Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language & Hearing Research*, 46(5), 1029–1037.
- Dawson, J., Stout, C., & Eyer, J. (2003). *Structured photographic expressive language test, third edition*. DeKalb, IL: Janelle.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.
- Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore: Paul H. Brookes Pub.
- Dorothy Bishop, & Becky Clark. (2012, October 24). Spotting specific language impairment: Signs for teachers to look out for. Retrieved from <http://www.theguardian.com/teacher-network/teacher-blog/2012/oct/24/specific-language-impairment-pupil-difficulties>
- Dunn, L. M., & Dunn, D. (2012). Peabody Picture Vocabulary Test, (PPVT-4). *Johannesburg: Pearson Education Inc.*
- Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact diagnostic decisions? *Child Language Teaching & Therapy*, 26(1), 77–92.
- Grist, M., & Hartshorne, M. V. (2014). Please! No more spaghetti...Commentary on Reilly et al., 2014. Specific language impairment: A convenient label for whom? *International Journal of Language & Communication Disorders*, 49(4), 416–451. <http://doi.org/10.1111/1460-6984.12102>
- Hammill, D. D., & Larsen, S. C. (2009). *Test of written language-Fourth edition*. Austin, TX: Pro-Ed.
- Hammill, D. D., & Newcomer, P. L. (2008). *Test of language development - Intermediate, fourth edition* (fourth). Austin, TX: PRO-ED.
- Harbour, R., & Miller, J. (2001). A new system for grading recommendations in evidence based guidelines. *BMJ : British Medical Journal*, 323(7308), 334–336.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- 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(3), 266–280.  
[http://doi.org/10.1044/0161-1461\(2013/12-0041\)](http://doi.org/10.1044/0161-1461(2013/12-0041))
- Hoffman, L. M., Loeb, D. F., Brandel, J., & Gillam, R. B. (2011). Concurrent and construct validity of oral language measures with school-age children with specific language impairment. *Journal of Speech, Language & Hearing Research*, 54(6), 1597–1608. [http://doi.org/10.1044/1092-4388\(2011/10-0213\)](http://doi.org/10.1044/1092-4388(2011/10-0213))
- Julie Dockrell, & Geoff Lindsay. (2014). The view from the chalk face. Commentary on Reilly et al., 2014. Specific language impairment: A convenient label for whom? *International Journal of Language & Communication Disorders*, 49(4), 416–451.  
<http://doi.org/10.1111/1460-6984.12102>
- Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 53(3), 739–756. [http://doi.org/10.1044/1092-4388\(2009/08-0248\)](http://doi.org/10.1044/1092-4388(2009/08-0248))
- Koutsoftas, A. D., & Gray, S. (2012). Comparison of narrative and expository writing in students with and without language-learning disabilities. *Language, Speech, and Hearing Services in Schools*, 43(4), 395–409. [http://doi.org/10.1044/0161-1461\(2012/11-0018\)](http://doi.org/10.1044/0161-1461(2012/11-0018))
- Leonard, L. B. (1991). Specific language impairment as a clinical category. *Language, Speech, and Hearing Services in Schools*, 22(2), 66–68.
- Leonard, L. B. (1995). Functional categories in the grammars of children with specific language impairment. *Journal of Speech & Hearing Research*, 38(6), 1270–1283.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Leonard, L. B. (2014). *Children with specific language impairment* (Second edition).

Cambridge, Massachusetts ; London, England: The MIT Press.

Mabel L. Rice. (2014). Advocating for SLI. Commentary on Bishop, D.V.M., 2014. Ten

questions about terminology for children with unexplained language problems.

*International Journal of Language and Communication Disorders*, (49), 381–415.

<http://doi.org/10.1111/1460-6984.12101>

McCauley, R. J., & Swisher, L. (1984). Psychometric review of language and articulation

tests for preschool children. *The Journal Of Speech And Hearing Disorders*,

49(1), 34–42.

McGoey, K., Cowan, R., Rumrill, P., & LaVogue, C. (2010). Understanding the

psychometric properties of reliability and validity in assessment. *Work*, 36(1),

105–111. <http://doi.org/10.3233/WOR-2010-1012>

Meline, T., & Paradiso, T. (2003). Evidence-based practice in schools: Evaluating

research and reducing barriers. *Language, Speech, and Hearing Services in*

*Schools*, 34(4), 273–283.

Montgomery, J. K. (2008). *Montgomery assessment of vocabulary acquisition*.

Greenville, SC: Super Duper.

Newcomer, P. L., & Hammill, D. D. (2008). *Test of language development - Primary*,

*fourth edition* (fourth). Austin, TX: PRO-ED.

Newcomer, P.S., & Hammill, D. D. (1997). *Test of language development-Primary, fourth*

*edition*. Austin, TX: PRO-ED.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- Nippold, M. A., Mansfield, T. C., Billow, J. L., & Tomblin, J. B. (2008). Expository discourse in adolescents with language impairments: Examining syntactic development. *American Journal of Speech-Language Pathology*, 17(4), 356–366.
- Nippold, M. A., Mansfield, T. C., Billow, J. L., & Tomblin, J. B. (2009). Syntactic development in adolescents with a history of language impairments: A follow-up Investigation. *American Journal of Speech-Language Pathology*, 18(3), 241–251.
- 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.
- 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.  
<http://doi.org/10.1044/0161-1461.2501.15>
- 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.  
<http://doi.org/10.1111/1460-6984.12102>
- Rice, M., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research*, 39(6), 1239–1257.
- Rice, M., & Wexler, K. (2001). *Rice/Wexler test of early grammar impairment*. San Antonio, TX: Psychological Corp.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- Rice, M., Wexler, K., Marquis, J., & Hershberger, S. (2000). Acquisition of irregular past tense by children with specific language impairment. *Journal of Speech, Language & Hearing Research*, 43(5), 1126–1145.
- Rutter, M. (2014). Classificatory purposes and diagnostic concepts. Commentary on Reilly et al., 2014. Specific language impairment: A convenient label for whom? *International Journal of Language & Communication Disorders*, 49(4), 416–451. <http://doi.org/10.1111/1460-6984.12102>
- Schuele, C. M., & Dykes, J. C. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment. *Clinical Linguistics & Phonetics*, 19(4), 295–318.
- Scott, C. M., & Balthazar, C. H. (2010). The grammar of information: Challenges for older students with language impairments. *Topics in Language Disorders*, 30(4), 288–307.
- Scott, C. M., & Stokes, S. L. (1995). Measures of syntax in school-age children and adolescents. *Language, Speech, and Hearing Services in Schools*, 26(4), 309–319.
- Semel, E., Wiig, E. H., & Secord, W. (2003). *Clinical evaluation of language fundamentals, fourth edition* (fourth). San Antonio, TX: Pearson.
- Semel, E., Wiig, E. H., & Secord, W. (2004). *Clinical evaluation of language fundamentals - Preschool-2* (Second). San Antonio, TX: Pearson.
- Semel, E., Wiig, E. H., & Secord, W. (2013). *Clinical evaluation of language fundamentals, fifth edition* (fifth). San Antonio, TX: Pearson.



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- Shriberg, L., Tomblin, J., & McSweeney, J. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language & Hearing Research*, 42(6), 1461–1481.
- Spaulding, T., Plante, E., & Farinella, K. (2006). Eligibility criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech, and Hearing Services in Schools*, 37(1), 61–72.
- St Clair, M. C., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2011). A longitudinal study of behavioral, emotional and social difficulties in individuals with a history of specific language impairment (SLI). *Journal of Communication Disorders*, 44(2), 186–199. <http://doi.org/10.1016/j.jcomdis.2010.09.004>
- Stephen Parsons, Emma Jordan, & Anna Branagan. (2014). Changes to specific language impairment: The service perspective. Commentary on Reilly et al., 2014. Specific language impairment: A convenient label for whom? *International Journal of Language & Communication Disorders*, 49(4), 416–451.  
<http://doi.org/10.1111/1460-6984.12102>
- Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics*. Boston: Pearson Education..
- 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(4), 236–41.
- Windsor, J., Scott, C. M., & Street, C. K. (2000). Verb and noun morphology in the spoken and written language of children with language learning disabilities.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

*Journal of Speech, Language, and Hearing Research*, 43(6), 1322–1336.

<http://doi.org/10.1044/jslhr.4306.1322>

Zipoli, R., Jr., & Kennedy, M. (2005). Evidence-based practice among speech-language pathologists: Attitudes, utilization, and barriers. *American Journal of Speech-Language Pathology*, 14(3), 208–220.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### Appendix A. Questionnaire

*Selection includes initial, elementary grade, and demographics portions*

Thank you for participating in this survey! Before you continue, if you have any questions about this survey, please contact Diane Ogiela or Jennifer Montzka at 208-373-1870 or [langlab@isu.edu](mailto:langlab@isu.edu).

#### Consent to Participate in This Study

I have read and understand the information provided in the study invitation posting or email. I understand that I may discontinue my participation in this study at any time. I have been given an opportunity to ask questions by emailing the investigator and, if I have contacted the investigator, all of my questions have been answered to my satisfaction.

1. By answering “Yes,” I willingly agree to participate in the research previously described.

- ☐ Yes, I agree to voluntarily participate in this study.
- ☐ No, I do not agree to participate

2. Please select the most accurate answer below.

- ☐ I am not a speech-language pathologist.
- ☐ I am a speech-language pathologist who DOES NOT work in a school setting.
- ☐ I am a speech-language pathology assistant or aide who works in a school setting.
- ☐ I am a speech-language pathologist who works at least partially in a school setting without licensure or certification.
- ☐ I am a speech-language pathologist who works at least partially in a school setting with temporary/conditional licensure or certification.
- ☐ I am a licensed or certified speech-language pathologist who works at least partially in a school setting.

3. How much of your caseload/workload do you estimate takes place in a school setting?

- ☐ 0-20%
- ☐ 21-40%
- ☐ 41-60%
- ☐ 61-80%
- ☐ 81-100%

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

4. The majority of children on your caseload are in:
- ☐ Preschool and Kindergarten
  - ☐ Elementary school (1st through 5th grade)
  - ☐ Middle and/or High school (6th through 12th grade)
5. Are at least 30% of the children on your caseload in preschool and/or kindergarten?
- ☐ Yes
  - ☐ No
6. Are at least 30% of the children on your caseload in elementary school (1st to 5th grade)?
- ☐ Yes
  - ☐ No
7. Are at least 30% of the children on your caseload in middle and high school (6th to 12th grade)?
- ☐ Yes
  - ☐ No

Please consider the following scenario when answering the survey questions:

You are preparing to evaluate or re-evaluate an elementary school (1st - 5th grade) child who is suspected of having a language impairment. The child is a native speaker of English with typical hearing and typical nonverbal cognitive skills, has no diagnosed developmental disorders or syndromes (e.g. Down Syndrome, Fragile X Syndrome, Autism Spectrum Disorder, etc.), has no history of brain injury/insult, and oral language is his/her primary mode of communication.

How likely are you to use the following types of assessment tools to evaluate the child's language skills?

8. A non-standardized assessment measure that evaluates many areas of language, such as a written or spoken language sample analysis.
- ☐ Very Likely
  - ☐ Likely
  - ☐ Neither likely nor unlikely
  - ☐ Unlikely
  - ☐ Very Unlikely

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

9. A non-standardized assessment measure that evaluates a specific area of language, such as clinician-generated probes for morpheme use or syntactic structures, comprehension of questions, pragmatic functions, content-specific vocabulary, etc.

- ☐ Very Likely
- ☐ Likely
- ☐ Neither likely nor unlikely
- ☐ Unlikely
- ☐ Very Unlikely

10. A standardized norm-referenced test that uses multiple subtests to evaluate many areas of receptive and/or expressive language.

- ☐ Very Likely
- ☐ Likely
- ☐ Neither likely nor unlikely
- ☐ Unlikely
- ☐ Very Unlikely

11. A standardized norm-referenced test that evaluates a specific area of language (morphology, syntax, semantics, pragmatics).

- ☐ Very Likely
- ☐ Likely
- ☐ Neither likely nor unlikely
- ☐ Unlikely
- ☐ Very Unlikely

12. When you use an assessment tool (standardized or non-standardized) that evaluates a specific area of language, you are most likely to administer it...

- ☐ Instead of administering a tool that evaluates many areas of language
- ☐ In addition to administering a tool that evaluates many areas of language

13. When you use an assessment tool that evaluates a specific area of language in addition to administering one that evaluates many areas of language, you are most likely to administer it...

- ☐ In order to add information about different areas not addressed by the initial assessment
- ☐ In order to add more detailed information about areas already addressed by the initial assessment

The following questions will all relate to standardized, norm-referenced language tests.

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Please consider the following scenario when answering the survey questions:

You are preparing to evaluate or re-evaluate an elementary school (1st - 5th grade) child who is suspected of having a language impairment. The child is a native speaker of English with typical hearing and typical nonverbal cognitive skills, has no diagnosed developmental disorders or syndromes (e.g. Down Syndrome, Fragile X Syndrome, Autism Spectrum Disorder, etc.), has no history of brain injury/insult, and oral language is his/her primary mode of communication.

14. If you received the referral above, which of the following language tests would you be most likely to administer to this child? Please select only 1 test; you will have an opportunity to select additional tests next. Please scroll down to see the list of all available tests before making your choice.

- ☐ I would not administer any standardized language tests
- ☐ Battelle Developmental Inventory, Second Edition (BDI-2)
- ☐ Children's Communication Checklist-2 (CCC-2)
- ☐ Clinical Evaluation of Language Fundamentals - Fifth Edition (CELF-5)
- ☐ Clinical Evaluation of Language Fundamentals - Fourth Edition (CELF-4)
- ☐ Communication Abilities Diagnostic Test (CADet)
- ☐ Comprehensive Assessment of Spoken Language (CASL)
- ☐ Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)
- ☐ Comprehensive Test of Phonological Processing (CTOPP)
- ☐ Diagnostic Evaluation of Language Variation-Norm-Referenced (DELV-N)
- ☐ Expressive One-Word Picture Vocabulary Test, Fourth Edition (EOWPVT-4)
- ☐ Expressive One-Word Picture Vocabulary Test, Third Edition (EOWPVT-3)
- ☐ Expressive Vocabulary Test, Second Edition (EVT-2)
- ☐ HELP Test-Elementary (HELP)
- ☐ Illinois Test of Psycholinguistic Abilities, Third Edition (ITPA-3)
- ☐ Language Processing Test 3 Elementary (LPT-3)
- ☐ Montgomery Assessment of Vocabulary Acquisition (MAVA)
- ☐ Oral and Written Language Scales, Second Edition (OWLS-II)
- ☐ Oral and Written Language Scales (OWLS)
- ☐ Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)
- ☐ Receptive One-Word Picture Vocabulary Test, Fourth Edition (ROWPVT-4)
- ☐ Receptive One-Word Picture Vocabulary Test, Third Edition (ROWPVT-3)
- ☐ Rice Wexler Test of Early Grammatical Impairment (TEGI)
- ☐ Structured Photographic Expressive Language Test 3 (SPELT-3)
- ☐ Test for Auditory Comprehension of Language, Fourth Edition (TACL-4)
- ☐ Test for Auditory Comprehension of Language, Third Edition (TACL-3)
- ☐ Test of Auditory Processing Skills - 3 (TAPS-3)

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- ☐ Test of Language Development-Intermediate, Fourth Edition (TOLD-I:4)
- ☐ Test of Language Development-Primary, Fourth Edition (TOLD-P:4)
- ☐ Test of Narrative Language (TNL)
- ☐ Test of Problem Solving 3 Elementary (TOPS-3)
- ☐ Test of Reception of Grammar - Version 2 (TROG-2)
- ☐ Test of Semantic Skills Primary (TOSS-P)
- ☐ Test of Word Finding in Discourse (TWF-D)
- ☐ Test of Word Finding, Third Edition (TWF-3)
- ☐ Test of Word Finding, Second Edition (TWF-2)
- ☐ Test of Word Knowledge (TOWK)
- ☐ Test of Written Language — Fourth Edition (TOWL-4)
- ☐ The Listening Comprehension Test - 2 (LCT-2)
- ☐ The WORD Test -3 Elementary (WORD-3E)
- ☐ The WORD Test -2 Elementary (WORD-2E)
- ☐ Token Test for Children, Second Edition (Token Test-2)
- ☐ OTHER \_\_\_\_\_

15. What additional tests would you be likely to administer to this an elementary school (1st - 5th grade) child as a supplement to the test you selected in the previous question? Choose up to 5 tests; please do not select the test you indicated in the previous question. Please scroll down to see the list of all available tests before making your choice.

- ☐ I would not administer additional standardized language tests
- ☐ Battelle Developmental Inventory, Second Edition (BDI-2)
- ☐ Children's Communication Checklist - 2 (CCC-2)
- ☐ Clinical Evaluation of Language Fundamentals - Fifth Edition (CELF-5)
- ☐ Clinical Evaluation of Language Fundamentals - Fourth Edition (CELF-4)
- ☐ Communication Abilities Diagnostic Test (CADet)
- ☐ Comprehensive Assessment of Spoken Language (CASL)
- ☐ Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)
- ☐ Comprehensive Test of Phonological Processing (CTOPP)
- ☐ Diagnostic Evaluation of Language Variation-Norm-Referenced (DELV-N)
- ☐ Expressive One-Word Picture Vocabulary Test, Fourth Edition (EOWPVT-4)
- ☐ Expressive One-Word Picture Vocabulary Test, Third Edition (EOWPVT-3)
- ☐ Expressive Vocabulary Test, Second Edition (EVT-2)
- ☐ HELP Test-Elementary (HELP)
- ☐ Illinois Test of Psycholinguistic Abilities, Third Edition (ITPA-3)
- ☐ Language Processing Test 3 Elementary (LPT-3)
- ☐ Montgomery Assessment of Vocabulary Acquisition (MAVA)

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- ☐ Oral and Written Language Scales, Second Edition (OWLS-II): Listening Comprehension / Oral Expression
- ☐ Oral and Written Language Scales, Second Edition (OWLS-II): Reading Comprehension / Written Expression
- ☐ Oral and Written Language Scales (OWLS): Listening Comprehension / Oral Expression
- ☐ Oral and Written Language Scales (OWLS): Reading Comprehension / Written Expression
- ☐ Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)
- ☐ Receptive One-Word Picture Vocabulary Test, Fourth Edition (ROWPVT-4)
- ☐ Receptive One-Word Picture Vocabulary Test, Third Edition (ROWPVT-3)
- ☐ Rice Wexler Test of Early Grammatical Impairment (TEGI)
- ☐ Structured Photographic Expressive Language Test 3 (SPELT-3)
- ☐ Test for Auditory Comprehension of Language, Fourth Edition (TACL-4)
- ☐ Test for Auditory Comprehension of Language, Third Edition (TACL-3)
- ☐ Test of Auditory Processing Skills 3 (TAPS-3)
- ☐ Test of Language Development-Intermediate, Fourth Edition (TOLD-I:4)
- ☐ Test of Language Development-Primary, Fourth Edition (TOLD-P:4)
- ☐ Test of Narrative Language (TNL)
- ☐ Test of Problem Solving 3 Elementary (TOPS-3)
- ☐ Test of Reception of Grammar - Version 2 (TROG-2)
- ☐ Test of Semantic Skills Primary (TOSS-P)
- ☐ Test of Word Finding in Discourse (TWF-D)
- ☐ Test of Word Finding, Third Edition (TWF-3)
- ☐ Test of Word Finding, Second Edition (TWF-2)
- ☐ Test of Word Knowledge (TOWK)
- ☐ Test of Written Language — Fourth Edition (TOWL-4)
- ☐ The Listening Comprehension Test 2 (LCT-2)
- ☐ The WORD Test 3 Elementary (WORD-3E)
- ☐ The WORD Test 2 Elementary (WORD-2E)
- ☐ Token Test for Children, Second Edition (Token Test-2)
- ☐ OTHER \_\_\_\_\_
- ☐ OTHER \_\_\_\_\_
- ☐ OTHER \_\_\_\_\_



## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

16. In your process of selecting a test, how much was your decision affected or unaffected by... Please scroll down to see all listed factors.

	Strongly affected	Slightly affected	Neither affected nor unaffected	Slightly unaffected	Strongly unaffected
...the time commitment required to administer (not score) the test?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...the time commitment required to score (not administer) the test?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...personal familiarity with the test based on past use?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...specific recommendations from other SLPs, graduate program faculty, or clinical supervisors?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...reviews of the test published in professional journals?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...publisher's description of the test?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...the cost of the test or test materials?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...the availability of the test at your school/facility?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...guidelines in your school/facility?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...guidelines in state and federal laws?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...psychometric features of the test (such as reliability and validity)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...specific measures of diagnostic accuracy (such as mean group differences or sensitivity and specificity) for the test as reported in the examiner's manual or in published research articles?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

17. Of the previously listed factors, select the factor you consider to be the most important in your decision to use a particular test:

- ☐ Time commitment to administer (not score) the test
- ☐ Time commitment to score (not administer) the test
- ☐ Personal familiarity with the test based on past use
- ☐ Specific recommendations from other SLPs, graduate program faculty, or clinical supervisors
- ☐ Reviews of the test published in professional journals
- ☐ Publisher's description of the test
- ☐ Cost of the test
- ☐ Availability of the test
- ☐ Facility regulations/restrictions
- ☐ State and Federal regulations/restrictions
- ☐ Psychometric features (such as reliability and validity) of the test
- ☐ Diagnostic accuracy (such as mean group differences or sensitivity and specificity) of the test
- ☐ None of the listed factors are the most important factor in the selection process

18. Because none of the listed factors are the most important in your decision, what do you consider to be the most important factor?

19. Of the previously listed factors, select the factor you consider to be the least important in your decision to use a particular test:

- ☐ Time commitment to administer (not score) the test
- ☐ Time commitment to score (not administer) the test
- ☐ Personal familiarity with the test based on past use
- ☐ Specific recommendations from other SLPs, graduate program faculty, or clinical supervisors
- ☐ Reviews of the test published in professional journals
- ☐ Publisher's description of the test
- ☐ Cost of the test
- ☐ Availability of the test
- ☐ Facility regulations/restrictions

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

- ☐ State and Federal regulations/restrictions
- ☐ Psychometric features (such as reliability and validity) of the test
- ☐ Diagnostic accuracy (such as mean group differences or sensitivity and specificity) of the test
- ☐ None of the listed factors are the least important factor in the selection process

20. Because none of the listed factors are the least important in your decision, what do you consider to be the least important factor?

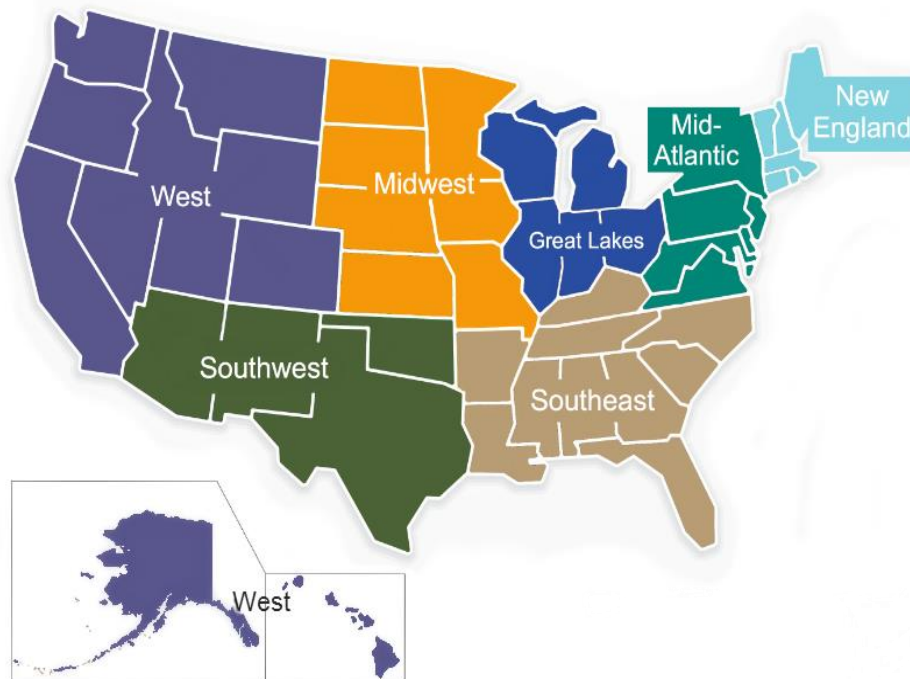
21. What is your primary resource in determining the overall psychometric quality (degree of reliability, validity, etc.) of a test?

- ☐ Other professionals' informal reports of the quality of the test
- ☐ Personal judgment based on prior experience with the test
- ☐ Reviews of the test published in professional journals
- ☐ The test's examiner's manual
- ☐ The publisher's catalog description of the test

22. How much of your school-based caseload/workload do you estimate occurs in each of the following grades?

	0 - 10%	11- 20%	21- 30%	31- 40%	41- 50%	51- 60%	61- 70%	71- 80%	81- 90%	91- 100%
Preschool and Kindergarten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1st through 5th grade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6th through 8th grade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
High school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS



23. In what region(s) of the United States have you worked as an SLP in a school-based setting? Please select all that apply.

- ☐ New England
- ☐ Mid-Atlantic
- ☐ Southeast
- ☐ Great Lakes
- ☐ Midwest
- ☐ Southwest
- ☐ West

24. In what region of the United States do you currently work as an SLP in a school-based setting? Please select only one.

- ☐ New England
- ☐ Mid-Atlantic
- ☐ Southeast
- ☐ Great Lakes
- ☐ Midwest
- ☐ Southwest
- ☐ West

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

25. How many years have you worked as an SLP in a school-based setting?
- ☐ 0-5
  - ☐ 6-10
  - ☐ 10-15
  - ☐ 16-20
  - ☐ 21-25
  - ☐ 25+
26. What is your highest level of education?
- ☐ Bachelor's degree
  - ☐ Master's degree
  - ☐ Doctorate (either PhD or SLPD)
27. What certifications and/or licenses do you currently hold? Please select all that apply.
- ☐ Certificate of Clinical Competence issued by the American Speech-Language Hearing Association
  - ☐ State-issued health provider's license
  - ☐ State-issued educational provider's license
  - ☐ State-issued working waiver and/or temporary/conditional license

**THANK YOU**  
for participating in this survey!  
Your responses have been recorded.

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

## Appendix B. List of Tests Provided to Participants

Test	Reference	Mini- mum Age	Maxi- mum Age
Battelle Developmental Inventory, Second Edition (BDI-2)	Newborg, J. (2004). <i>Battelle developmental inventory, second edition</i> . Rolling Meadows, IL: Riverside.	0	8
Clinical Evaluation of Language Fundamentals--Fifth Edition (CELF-5)	Semel, E., Wiig, E. H., & Secord, W. (2013). <i>Comprehensive evaluation of language fundamentals, fifth edition</i> . San Antonio, TX: Pearson.	5	21
Clinical Evaluation of Language Fundamentals--Fourth Edition (CELF-4)	Semel, E., Wiig, E. H., & Secord, W. (2003). <i>Comprehensive evaluation of language fundamentals, fourth edition</i> . San Antonio, TX: Pearson.	5	21
Communication Abilities Diagnostic Test (CADet)	Johnston, E. B., & Johnston, A. V. (1990). <i>Communication abilities diagnostic test</i> . Austin, TX: PRO-ED.	3	9
Comprehensive Assessment of Spoken Language (CASL)	Carrow-Woolfolk, E. (1999). <i>Comprehensive assessment of spoken language</i> . Circle Pines, MN: Western Psychological Services.	3	21
Comprehensive Test of Phonological Processing – Second Edition (CTOPP-2)	Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). <i>Comprehensive test of phonological processing, second edition</i> . Austin, TX: PRO-ED.	4	24
Comprehensive Test of Phonological Processing (CTOPP)	Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). <i>Comprehensive test of phonological processing</i> . Austin, TX: PRO-ED.	4	24
Diagnostic Evaluation of Language Variation-Norm- Referenced (DELV-N)	Seymour, H. N., Roeper, T. W., & de Villiers, J. (2005). <i>Diagnostic evaluation of language variation--Norm-referenced</i> . San Antonio, TX: Pearson.	4	9
Expressive One-Word Picture Vocabulary Test (EOWPVT-3)	Brownell, R. (2000). <i>Expressive one-word picture vocabulary test, third edition</i> . Novato, CA: Academic Therapy.	2;0	80+
Expressive One-Word Picture Vocabulary Test (EOWPVT-4)	Brownell, R. (2011). <i>Expressive one-word picture vocabulary test, fourth edition</i> . Novato, CA: Academic Therapy.	2;0	80+

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Test	Reference	Mini- mum Age	Maxi- mum Age
Expressive Vocabulary Test, Second Edition (EVT-2)	Williams, K. T. (2007). <i>Expressive vocabulary test, second edition</i> . San Antonio, TX: Pearson.	2;5	90
HELP Test-Elementary (HELP)	Lazzari, A. M., & Peters, P. M. (1993). <i>HELP -Elementary</i> . East Moline, IL: LinguiSystems.	6;0	11;11
Illinois Test of Psycholinguistic Abilities-Third Edition (ITPA-3)	Hammill, D. D., Mather, N., & Roberts, R. (2001). <i>Illinois test of psycholinguistic abilities, third edition</i> . Austin, TX: PRO-ED.	5;0	11;11
Language Processing Test 3 Elementary (LPT-3)	Richard, G. J., & Hanner, M. A. (2005). <i>Language processing test 3: Elementary</i> . East Moline, IL: LinguiSystems.	5;0	11;11
Montgomery Assessment of Vocabulary Acquisition (MAVA)	Montgomery, J. K. (2008). <i>Montgomery assessment of vocabulary acquisition</i> . Greenville, SC: Super Duper.	3;0	12;11
Oral and Written Language Scales (OWLS)	Carrow-Woolfolk, E. (1996). <i>Oral and written language scales</i> . Torrance, CA: Western Psychological Services.	3;0	20;11
Oral and Written Language Scales, Second Edition (OWLS-II)	Carrow-Woolfolk, E. (2011). <i>Oral and written language scales, second edition</i> . Torrance, CA: Western Psychological Services.	3;0	20;11
Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)	Dunn, L. M., & Dunn, D. M. (2007). <i>Peabody picture vocabulary test, fourth edition</i> . San Antonio, TX: Pearson.	2;6	90+
Receptive One-Word Picture Vocabulary Test, Fourth Edition (ROWPVT-4)	Brownell, R. (Ed.). (2010). <i>Receptive one-word picture vocabulary test</i> . San Antonio, TX: Pearson.	2;0	19;11
Receptive One-Word Picture Vocabulary Test, Third Edition (ROWPVT-3)	Brownell, R. (2000b). <i>Receptive one-word picture vocabulary test, third edition</i> . Novato, CA: Academic Therapy.	2;0	80+
Rice Wexler Test of Early Grammatical Impairment (TEGI)	Rice, M., & Wexler, K. (2001). <i>Rice/Wexler test of early grammar impairment</i> . San Antonio, TX: Psychological Corp.	3;0	7;11
Structured	Dawson, J., Stout, C., & Eyer, J. (2003).	4;0	8;11

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Test	Reference	Mini- mum Age	Maxi- mum Age
Photographic Expressive Language Test 3 (SPELT-3)	<i>Structured photographic expressive language test, third edition.</i> DeKalb, IL: Janelle.		
Test for Auditory Comprehension of Language-3rd Edition (TACL-3)	Carrow-Woolfolk, E. (1999b). <i>Test for auditory comprehension language, third edition.</i> Austin, TX: PRO-ED.	3;0	8;11
Test for Auditory Comprehension of Language-4th Edition (TACL-4)	Carrow-Woolfolk, E. (2014). <i>Test for auditory comprehension of language, fourth edition.</i> Austin, TX: PRO-ED.	3;0	8;11
Test of Auditory Processing Skills-3 (TAPS-3)	Martin, N., & Brownell, R. (2005). <i>Test of auditory processing skills, third edition.</i> Novato, CA: Academic Therapy.	4;0	17;11
Test of Language Development-Intermediate, Fourth Edition (TOLD-I:4)	Hammill, D. D., & Newcomer, P. L. (2008). <i>Test of language development - Intermediate, fourth edition.</i> Austin, TX: PRO-ED.	8;0	16;11
Test of Language Development-Primary, Fourth Edition (TOLD-P:4)	Newcomer, P. L., & Hammill, D. D. (2008). <i>Test of language development - Primary, fourth edition.</i> Austin, TX: PRO-ED.	4;0	7;11
Test of Narrative Language (TNL)	Gillam, R. B., & Pearson, N. A. (2004). <i>Test of narrative language.</i> Austin, TX: PRO-ED.	5;0	10;11
Test of Problem Solving 3 Elementary (TOPS-3)	Bowers, L., Huisinigh, R., & LoGuidice, C. (2005). <i>Test of problem solving 3 - Elementary.</i> East Moline, IL: LinguiSystems.	6;0	11;11
Test of Reception of Grammar - Version 2 (TROG-2)	Bishop, D. (2003). <i>Test for reception of grammar, second edition.</i> San Antonio, TX: Pearson.	4;0	80+
Test of Semantic Skills Primary (TOSS-P)	Bowers, L., Huisinigh, R., LoGuidice, C., & Orman, J. (2002). <i>Test of semantic skills - Primary.</i> East Moline, IL: LinguiSystems.	4;0	7;11
Test of Word Finding in Discourse (TWF-D)	German, D. J. (1991). <i>Test of word finding in discourse.</i> Austin, TX: PRO-ED.	6;6	11;11
Test of Word Finding, Second Edition (TWF-2)	German, D. J. (2000). <i>Test of word finding—2.</i> East Moline, IL: LinguiSystems.	4;0	11;11



# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Test	Reference	Mini- mum Age	Maxi- mum Age
Test of Word Finding, Third Edition (TWF-3)	German, D. J. (2015). <i>Test of word finding—3</i> . East Moline, IL: LinguiSystems.	4;0	11;11
Test of Word Knowledge (TOWK)	Wiig, E., & Secord, W. (1991). <i>Test of word knowledge</i> . San Antonio, TX: Harcourt Assessment.	5;0	16;11
Test of Written Language — Fourth Edition (TOWL-4)	Hammill, D. D., & Larsen, S. C. (2009). <i>Test of written language-Fourth edition</i> . Austin, TX: Pro-Ed.	9;0	16;11
The Listening Comprehension Test-2 (LCT-2)	Huisingh, R., Bowers, L., & LoGiudice, C. (2006). <i>The listening comprehension test—2</i> . East Moline, IL: LinguiSystems.	6;0	10;11
The WORD Test 2 Elementary (WORD- 2E)	Bowers, L., Huisingh, R., LoGiudice, C., & Orman, J. (2004). <i>The WORD test 2 elementary</i> . East Moline, IL: LinguiSystems, Inc.	6;0	10;11
The WORD Test 3 Elementary (WORD- 3E)	Bowers, L., Huisingh, R., LoGiudice, C., & Orman, J. (2014). <i>The WORD test 3 elementary</i> . East Moline, IL: LinguiSystems, Inc.	6;0	10;11
Token Test for Children-2nd Edition (Token Test-2)	McGhee, R. L., Ehrler, D. J., & DiSimoni, F. (2007). <i>Token test for children—2nd edition</i> . Austin, TX: Pro-Ed.	3;0	11;11

All tests have minimum age no greater than 9 and were published in 1990 or later. Any test with a new edition published within the last five years has both the most recent edition and the previous edition.

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

## Appendix C. Additional Assessments Recommended

Test	Reference	Number of Recommendations
Boehm Test of Basic Concepts-3 (Boehm-3)	Boehm, A. E. (2000). <i>Boehm test of basic concepts, third edition</i> . San Antonio, TX: Pearson.	1
Bracken Basic Concept Scales	Bracken, B. A. (1998). <i>Bracken basic concept scale—Revised</i> . San Antonio, TX: The Psychological Corporation.	1
Clinical Assessment of Articulation and Phonology (CAAP)	Secord, W., Donohue, J., & Johnson, C. (2002). <i>Clinical assessment of articulation and phonology</i> . Greenville, SC: Super Duper Publications, Inc.	1
Clinical Evaluation of Language Fundamentals 4 (CELF-4) Phonological Awareness Criterion-Reference subtest	Semel, E., Wiig, E. H., & Secord, W. (2003). <i>Comprehensive evaluation of language fundamentals, fourth edition</i> . San Antonio, TX: Pearson.	1
Comprehensive Receptive and Expressive Vocabulary Test 2 (CREVT-2)	Wallace, G., & Hammill, D. D. (2002). <i>Comprehensive receptive and expressive vocabulary test</i> . Austin, TX: PRO-ED.	2
Comprehensive Receptive and Expressive Vocabulary Test 3 (CREVT-3)	Wallace, G., & Hammill, D. D. (2013). <i>Comprehensive receptive and expressive vocabulary test, third edition</i> . Austin, TX: PRO-ED.	2
Expressive Language Test (ELT)	Bowers, L., LoGuidice, C., Orman, J., & Huisinigh, R. (1998). <i>The expressive language test</i> . East Moline, IL: LinguiSystems.	2
Expressive Language Test 2 (ELT 2)	Bowers, L., Huisinigh, R., LoGuidice, C., & Orman, J. (2010). <i>The expressive language test, second edition</i> . East Moline, IL: LinguiSystems.	1
HearBuilder Phonological Awareness Test (H-PAT)	Wiig, E. H., & Secord, W. (2011). <i>HearBuilder phonological awareness test</i> . Greenville, SC: Super Duper.	1
Lindamood Auditory Conceptualization Test — Third Edition (LAC-3)	Lindamood, P. C., & Lindamood, P. (2004). <i>Lindamood auditory conceptualization test</i> . Austin, TX: PRO-ED.	1

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Test	Reference	Number of Recommendations
Phonological Awareness Test 2 (PAT-2)	Robertson, C., & Salter, W. (2007). <i>Phonological awareness test, second edition</i> . East Moline, IL: LinguiSystems.	3
Pragmatic Language Skills Inventory (PLSI)	Gilliam, J. A., & Miller, L. (2006). <i>Pragmatic language skills inventory</i> . Austin, TX: PRO-ED.	2
Preschool Language Scales 5 (PLS 5)	Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). <i>Preschool language scale, fifth edition, English edition</i> . San Antonio, TX: Pearson	1
Qualitative Reading Inventory 5 (QRI-5)	Leslie, L., & Schudt Caldwell, J. (2010). <i>Qualitative reading inventory 5</i> . San Antonio, TX: Pearson.	1
Social Language Development Test - Elementary (SLT-E)	Bowers, L., Huisinigh, R., & LoGuidice, C. (2008). <i>Social language development test: Elementary</i> . East Moline, IL: LinguiSystems.	6
Structured Photographic Expressive Language Test (SPELT)	Citation unavailable (1974).	1
Test for Examining Expressive Morphology (TEEM)	Shipley, K. G., Sue, M. B., & Stone, T. A. (1983). <i>Test for Examining Expressive Morphology</i> . Tuscon, AZ: Communication Skill Builders.	3
Test of Auditory Reasoning and Processing Skills (TARPS)	Gardner, M. F. (1993). <i>Test of auditory reasoning and processing skills</i> . Novato, CA: Academic Therapy.	1
Test of Language Competence (TLC)	Wiig, E. H., & Secord, W. (1989). <i>Test of language competence</i> . San Antonio, TX: Psychological Corporation.	1
Test of Pragmatic Language (TOPL)	Phelps-Terasaki, D., & Phelps-Gunn, T. (1992). <i>Test of pragmatic language</i> . San Antonio, TX: Psychological Corporation.	1
Test of Pragmatic Language-2 (TOPL-2)	Phelps-Terasaki, D., & Phelps-Gunn, T. (2007). <i>Test of pragmatic language, second edition</i> . Austin, TX: PRO-ED.	1
The Strong Narrative Assessment Procedure (SNAP)	Strong, C. J. (1998). <i>The Strong Narrative Assessment Procedure</i> . Austin, TX: PRO-ED.	1

# SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

## Appendix D. States Listed by Regions

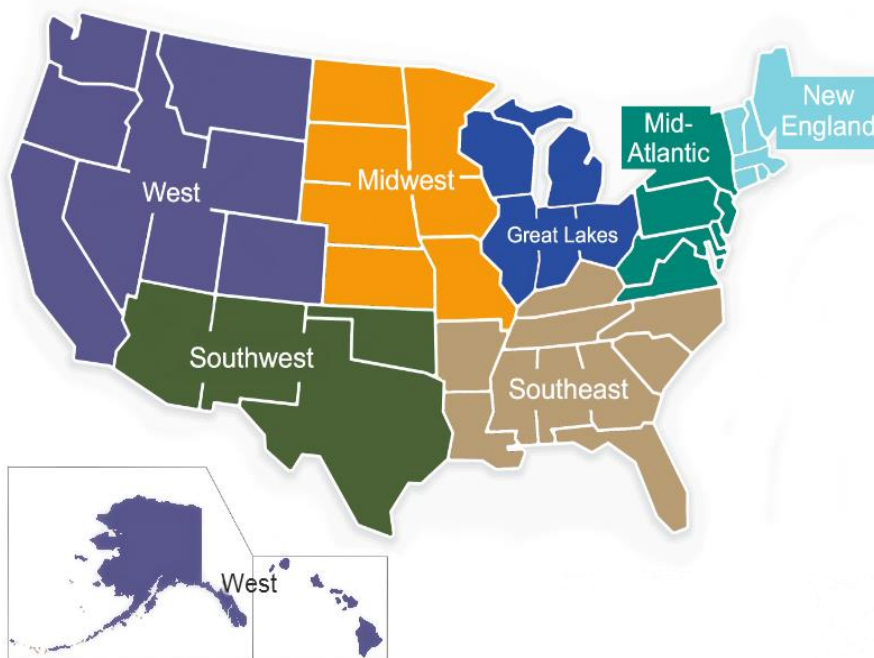


Figure B1. Map provided to questionnaire participants for demographics questions

Table B1

### *States Listed by Regions*

<b>New England</b>	<b>Great Lakes</b>	<b>Midwest</b>	<b>West</b>
Connecticut	Illinois	Iowa	Alaska
Maine	Indiana	Kansas	California
Massachusetts	Michigan	Minnesota	Colorado
New Hampshire	Ohio	Missouri	Hawaii
Rhode Island	Wisconsin	Nebraska	Idaho
Vermont		North Dakota	Montana
		South Dakota	Nevada
	<b>Southeast</b>		Oregon
<b>Mid-Atlantic</b>	Alabama		Utah
Delaware	Arkansas	<b>Southwest</b>	Washington
Maryland	Florida	Arizona	Wyoming
New Jersey	Georgia	New Mexico	
New York	Kentucky	Oklahoma	
Pennsylvania	Louisiana	Texas	
Virginia	Mississippi		
West Virginia	North Carolina		
	South Carolina		
	Tennessee		

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

### Appendix E. Participant-generated Most Important and Least Important Factors

Table E1

#### *Participant-Generated Most Important Factors*

Additional Considerations Listed by Participants in Response to “Most Important Factor” Question	Number of Additional Considerations	Descriptive Category of Additional Considerations
A strong baseline assessment that would give a true, overall view of the student's abilities based on the described deficits.	One	Undefined "High Quality"
The areas the assessment tests and how it tests those areas.	One	Assessment Components and Approach
Does it assess what I suspect the problem is and will I gain usable information for planning interventions.	Two	Targets Individual Deficits Relationship to Intervention
The concern regarding the student - what test would best measure the teacher's concerns.	Two	Targets Individual Deficits Targets Teacher Concerns
selecting a test that addresses the specific area(s) of difficulty that the student is having	One	Targets Individual Deficits
How fair it would be for the population I serve. The CELF-5 would be too hard because too many of our students have minimal life experiences.	One	Demographically Representative
Does the test specifically target what I believe to be the student's weakness? I need in-depth testing to probe what I believe is the student's weaknesses.	One	Targets Individual Deficits
The weaknesses noted on the teacher narrative and the child's reading ability. If reading ability is low, I try to minimize that requirement.	Two	Assessment Components and Approach Targets Teacher Concerns
The teacher's reason for referral and the reason they see language deficits.	One	Targets Teacher Concerns
Appropriateness of the test based on the child's needs	One	Targets Individual Deficits

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Additional Considerations Listed by Participants in Response to “Most Important Factor” Question	Number of Additional Considerations	Descriptive Category of Additional Considerations
What I believe would be best to use based on the information I have about the child.	One	Targets Individual Deficits
What was best suited for that individual student	One	Targets Individual Deficits
does it test what the teacher has concerns about re language	One	Targets Teacher Concerns
Does the test measure and relate to the difficulties the student is having	One	Targets Individual Deficits
how effective the test will assess language as well as the scope of information I will gather from the assessment	Two	Undefined "High Quality" Undefined Provides "Valuable" Information
Does this test measure and provide information to answer the questions I have about this student? If I want to decide eligibility, does it give the scores for a child this age. If I need to write a goal or objectives, does it provide information about the priority area I want to target.	Two	Targets Individual Deficits Relationship to Intervention
Does it assess the areas of concern that are reported for the child that I am evaluating.	One	Targets Individual Deficits
Specific concerns about this student.	One	Targets Individual Deficits
I choose the test based on reported areas of concern and which test can best measure that skill or skills.	One	Targets Individual Deficits
Selection of the test is based on student needs.	One	Targets Individual Deficits
does the test do what its says it does and can you plan for therapy from it.	Two	Listed Factor: Psychometric Features Relationship to Intervention
Reliability of test results that do not over or under identify language impairments especially when a full case study	One	Listed Factor: Psychometric Features

## SELECTING TESTS FOR CHILD LANGUAGE DISORDERS

Additional Considerations Listed by Participants in Response to “Most Important Factor” Question	Number of Additional Considerations	Descriptive Category of Additional Considerations
evaluation has been completed and you are comparing test results with other members of the diagnostic team.		Listed Factor: Diagnostic Accuracy
The availability of the test, how well the test measures the possible language issue	Two	Listed Factor: Availability Targets Individual Deficits
The test must assess how a child uses classroom language and not be heavily dependent on memory.	Two	Functional Assessment Assessment Components and Approach Targets Individual Deficits
Student based areas of need.	Two	Targets Individual Deficits

Table E2

### *Participant-Generated Least Important Factors*

Additional Considerations Listed by Participants in Response to “Least Important Factor” Question	Descriptive Category of Additional Considerations
I don't know	Unsure
time to score the test	Listed Factor
I think they are all important factors in choosing an appropriate test for an individual student.	All Important
Not sure	Unsure
Does not provide information that I deem to be as valuable to the assessment as chosen tests	Undefined "Less Valuable" information
having administrators and other professionals tell you what to use	Listed Factor
Time	Listed Factor
Recommendations from other SLP's	Listed Factor