VideoAssessmentTools

Language

Video Assessment Tool

(VAT)

TABLE OF CONTENTS

About the Author

Adriana Lavi, PhD, CCC-SLP is a licensed speech-language pathologist and a pioneer in the development of speech and language video-based assessment tools. She is the creator and author of the Clinical Assessment of Pragmatics (CAPs), as well as Video Assessment Tools, an online assessment platform that features the Articulation and Phonology Video Assessment Tool, the IMPACT Social Communication Rating Scale, the IMPACT Articulation and Phonology Rating Scale, etc. Additionally, Dr. Lavi is the creator of the Video Learning Squad, an online therapy platform that features the Social Squad and Stutter Squad.

For over a decade, Dr. Lavi owned Go2Consult where she supervised 35+ speech-language pathologists and clinical fellows across Southern California. Dr. Lavi has also served as an Assistant Professor at the Department of Communicative Disorders at Loma Linda University, and is the founder of the Lavi Institute for Research and Professional Development. She earned a master's degree in speech-language pathology from California State University at Sacramento and a PhD degree in Rehabilitation Sciences with an emphasis in speech-language pathology from Loma Linda University. Dr. Lavi was one of three students selected by the Bureau of Educational and Cultural Affairs of the US Department of State from the country of Moldova to study in the US in 2000. She has lived through and understands the culture of poverty. Her professional career has always focused on service delivery for students from low-income backgrounds. Dr. Lavi is the proud mother of four young, highly energetic boys.

Technical Publications Writer

Charlotte Granitto is a licensed speech-language pathologist and technical publications writer. She works as part of the Lavi Institute team and focuses on the research and development of video-based assessments and treatments for individuals with speech and language disorders. Charlotte also has experience working in a clinic, hospital, and school setting as a rehabilitation therapist and speech-language pathologist. She earned a Bachelor of Arts degree in Psychology with a focus in research from Wilfrid Laurier University and earned a Master's degree in Education specializing in speech-language pathology from the State University of New York at Buffalo State.

Acknowledgements

Sincere appreciation is extended to the families and children who generously offered their time and effort. Additional thanks goes to the expert panel and speech-language pathologists who field-tested the children.

Chapter

1

Overview of the Test

Language Video Assessment Tool Description

The *Language Video Assessment Tool (VAT)* is a norm-referenced language testing battery for children and young adults ages 6 through 18 years old. It is composed of four stand-alone tests. It is an accurate and reliable assessment tool that provides valid results on language comprehension, story retell, following directions, and morphology and sentence structure. Normative data of this test is based on a nationally representative sample of 1554 (typically developing) children and young adults in the United States.

Language Video Assessment Tool Areas

The test is composed of four independent and stand-alone tests: (a) Language Comprehension in Context, (b) Restating Information, (c) Following Directions, and (d) Morphology and Sentence Structure.

Testing Format

The Language VAT is composed of four tests. Each test is video-based and requires access to a laptop or tablet. Each test yields a standard score and percentile. While completing each test, students will be asked to complete trial items and then continue into test items.

Administration Time

Administration time for the test takes approximately 40-50 minutes.

Test Uses

The results of the *Language Video Assessment Tool* provide comprehensive information on language comprehension and oral expression development of children and young adults. It presents with four essential purposes:

a) To help identify developmental language disorders and determine eligibility for special education services (e.g., initial IEP based evaluations);
b) To help determine strengths and weakness within a variety of language domains (e.g., morphology, syntax, spoken language comprehension);

c) To help document progress in language skills, measure treatment efficacy or re-evaluate overall language profiles as part of triennial IEP based reviews;
d) To help analyze language skills in children and young adults for research purposes

Features

Unique Design of Using Video Based Scenes and Instructions

One of the most notable benefits of the *Language VAT* is its unique test design consisting of videos and pictures to engage students during testing. The videos are presented in relevant, life-like content, and the actors in the videos are from a wide variety of ethnic and cultural backgrounds. The videos are presented at a rate that is controlled for speed without being unnaturally slow. Vocabulary used in the videos is appropriate to the testing age range (5-0 through 18-0), and the real-life scenes are those which might be expected to occur in educational setting or real-life in general.

Language Video Assessment Tool Uses and Purpose

The information obtained from the *Language Video Assessment Tool* can help determine what areas a child has deficits in and how deficits in these areas may impact the child in both the classroom and in the home environment. The *Language VAT* should be used to evaluate children or young adults who have a suspected or previous diagnosis of a language disorder. This tool will assist in the identification or continued diagnosis of a spoken language comprehension and/or expressive language disorder. The results of the *Language VAT* provide clinicians information on children and young adult's ability to comprehend spoken language and use spoken language. By utilizing the *Language VAT*, we are able to develop a better understanding as to how a student's language abilities may impact their academic performance, progress in school, and social interactions.

Language Video Assessment Tool Area Descriptions

Language Comprehension in Context

The *Language Comprehension in Context* test evaluates a student's ability to attentively listen to a short passage and then answer questions about the presented story. This test requires students to listen carefully, understand the main idea, remember key details and the sequence of events, and to use critical thinking skills to interpret inferences and make predictions. Difficulties in language comprehension may impact a student across the school-age years when attempting to understand orally presented stories as well as when attempting to follow along with a classroom lesson, text, or video.

Sample Language Comprehension in Context Item: At the beginning of the test, the clinician explains the test to the examinee by saying, *"You are about to watch some videos about nature. Please watch and listen carefully. You will be allowed to watch each video only one time. I will ask you questions about each video. Please do your best to answer the questions. You can ask me to repeat the question. I will be able to repeat the question only one time. If you don't know or can't remember the answer, it's okay to say, 'I don't know' or 'I can't remember.' So please tell me now, what are you supposed to do?"* Then the clinician will play the video, and the examinee will be asked questions about what happened/what was discussed in the video clip.

Restating Information

The *Restating Information* test evaluates a student's ability to listen to, comprehend, and then retell a story immediately after the presentation of the story. Students are asked to listen carefully and then tell the story back to the clinician. Difficulties in listening to and retelling a story may impact a student across the school-age years in assignments that require organization, sequencing, or describing. Students who have difficulty with retelling a story may have difficulties with story comprehension and/or vocabulary.

Restating Information Item: At the beginning of the test, the clinician explains the test to the examinee by saying, *"You are going to watch some videos. Please watch and listen carefully. You will watch the video only one time. After you are done watching, I need you to tell me what the videos are about and all the details that you can remember. So, tell me, what are you supposed to do now?"*

Following Directions

The *Following Directions* test items evaluate an individual's ability to interpret spoken directions, recall those directions, and then act upon those spoken directions. A student's success on this test reflects his/her comprehension and immediate recall abilities. Difficulties in following and executing directions may impact a student across the school-age years when attempting to follow along with teacher instructions in the classroom and when completing school assignments that requires following procedural directives.

Sample Following Directions Item: At the beginning of the test, the clinician explains the test to the examinee by saying, *"I am going to show you some pictures and ask you to draw some things. First, I want you to listen while I ask you to do something. Then, after I say 'go,' you can follow the instructions and draw. I can repeat the instructions for you one time if you need me to. So, can you please tell me what you are supposed to do now?"* Trial A: Point to the bus. Go. Trial B: Draw a pumpkin. Go.

Morphology and Sentence Structure

The *Morphology and Sentence Structure* test evaluates a student's ability to use grammatical morphemes and to compose meaningful and grammatically correct sentences. This test requires students to describe what is going on in a picture, retrieve a word to complete a sentence, and/or formulate sentences based upon a picture and specific words. Difficulties in morphology and sentence structure may impact a student across the school-age years when telling a story, completing a writing assignment, or during discourse or conversation with peers.

Sample Morphology and Sentence Structure Item: At the beginning of the test, the clinician explains the test to the examinee by saying, *"You are going to watch some videos. After watching each video, you will be asked to answer a question about the video or you will be asked to complete a sentence about the video. Let's try one."*

Standardized Video-Based Assessment

In the past few decades, the speech and language pathology field has begun incorporating multimedia features such as audio and animation into interventions and assessment tools. Numerous studies have investigated the potential advantages of using multimedia on young children's language production. For example, Schlosser's et al. (2014) study found that children between three and five years old were able to name animated symbols more easily than static symbols. Another study found that children increased

their identification of action verbs when the target included animation (Mineo, Peischl, & Pennington, 2008). Verhallen, Bus, and De Jong (2006) investigated the positive effects of animated books on narrative comprehension and language skills. The researchers found that when 3 and 5 year old children, from low educated families, were presented with a story with static images versus a story with multimedia additions (e.g., cinematic techniques such as zoom, pans, and sound effects), the stories with multimedia additions provided a better framework for understanding stories and remembering linguistic information. Gazella and Stockman (2003) compared children's ability to retell a story after listening to a story with only audio versus listening to a story while seeing pictures (audiovisual). Gazella and Stockman (2003) found that for children 4 years of age, there was a difference observed in macrostructural aspects of children's narratives. Specifically, these children made more inferences and told better stories when presented with an audiovisual format. Diehm, Wood, Puhlman, and Callendar (2020) further examined the effect of story presentation (static pictures versus animated videos) on preschool children's narrative story retelling. The results of the story revealed that typically developing children demonstrated higher quantity and quality of language within a story retell in response to an animated video than after viewing static images. This study suggests that when presented with a short video versus a picture book, young children may produce longer narrative retells, use a wider variety of vocabulary, and produce more complex syntax (Diehm, Wood, Puhlman, & Callendar, 2020).

The incorporation of videos may also allow students to focus on the salient features of instructions, eliminating the distraction of human interactions such as sounds, body language, tone of voice, and smells. This may be particularly beneficial for students with autism spectrum disorder who may have difficulty with some human interactions that tend to be emotionally loaded (Neumann, 2004). With the use of videos, extraneous variables can also be controlled and/or eliminated and have students focus on a screen, which may minimize attentional and language requirements (Sherer, Pierce, Paredes, Kisacky, Ingersoll, & Schreibman, 2001).

Diehm, Wood, Puhlman, and Callendar (2020) investigated the effect of story presentation (static pictures versus animated videos) on preschool children's narrative story retelling. Typically developing children demonstrated higher quantity and quality of language within a story retell in response to an animated video than after viewing static images. This study suggests that when presented with a short video versus a picture book, young children may produce longer narrative retells, use a larger variety of vocabulary, and produce more complex syntax (Diehm, Wood, Puhlman, & Callendar, 2020).

Chapter

2

Theoretical Background of the Language Video Assessment Tool

Spoken language comprehension and oral expression, refers to the understanding and the use of spoken language across various contexts and social situations. Approximately 7% of children have deficits in language comprehension or language use and these difficulties can persist into the school-age years and interfere with communication, academics, and social interactions (Tomblin, Records, Buckwalter, Zhang, Smith, & O'Brien, 1997). Longitudinal studies have revealed that language impairments that persist into school age remain in adolescence (Conti-Ramsden & Durkin 2007) and adulthood (Johnson, Beitchman, & Brownlie, 1999; Clegg, Hollis, Mawhood, & Rutter, 2005), often with accompanying literacy deficits (Clegg, Hollis, Mawhood, & Rutter, 2005, Snowling & Hulme, 2000). Lindsay and Dockrell (2012) conducted a longitudinal study with adolescents who were identified as having specific language impairment (SLI) during the early primary grades. This study assessed the behavioral, emotional, and social difficulties of students in relation to self-concept, language, and literacy abilities over time. Participants were followed from 8 years old to 17 years old. Lindsay and Dockrell (2012) found that poor language and literacy skills continued, and peer and conduct problems were found to increase over this age range. Joffee and Black (2012) explored behavioral, emotional, and social difficulties in young adolescents who, based on teacher report, were identified as having low language skills and/or poor academic achievement. These students had not been clinically diagnosed as having a language disorder. Results of Joffee and Black's (2012) study indicate that even students with subtle language problems can negatively impact school and social interactions. The researchers emphasized the need to identify and treat language weakness in all students so that all children can be properly supported.

Previous research findings have suggested that nearly all school-based speech-language pathologists use standardized testing as part of their student's diagnostic tools (Caesar & Kohler, 2009), and that half of SLPs regard standardized testing as the most important data collected during their evaluations and all SLPs ranked standardized testing in the top five most important tools (Eickhoof, Betz, & Ristow, 2010). Huang, Hopkins, and Nippold (1997) found that school-based SLPs were unsatisfied with current standardized testing and this was related to the time it took to actually complete the assessments and the time workplaces gave SLPs to finish the assessments. Thus, there is a clear need for both formal and informal assessment tools that aid in the identification of language disorders that are reliable, valid, and also timesaving. Without appropriate assessment and intervention, there can be serious negative impacts to a child's development, education, and social interactions. For example, a child who has difficulty with their ability to understand spoken language may find it difficult to follow along during classroom

instruction and fall behind in their classwork. A child who has trouble understanding or using spoken language may have difficulty developing meaningful peer relationships and friendships, which could lead to a variety of other difficulties such as behavioral and emotional problems. By assessing students with the *Language VAT,* speech-language pathologists are able to evaluate students with multimedia features (e.g., audio, animation, music) and tools (e.g., computers, tablets) that align with children's interests during this technological age they are growing up in. SLPs can depend on this effective, reliable, and valid formalized assessment tool to evaluate student's abilities and identify those individuals who have a suspected or an existing diagnosis of a language disorder and the impact the language disorder may have on the child.

Contextual Background for Test Areas

Language impairment involves difficulty in the understanding and/or use of spoken, written, and/or other symbol systems. The disorder may involve: "(1) the form of language (phonology, morphology, syntax); (2) the content of language (semantics); and/or (3) the function of language in communication (pragmatics) in any combination" (ASHA, 2016). Listening comprehension is a high-order skill that involves both language and cognitive abilities (Florit, Roch, & Levorato, 2013; Kim & Phillips, 2014; Lepola, Lynch, Laakkonen, Silven, & Niemi, 2012). Specifically, listening comprehension refers to one's ability to comprehend spoken language (e.g., conversations, stories/narratives) by extracting and constructing meaning. Research has showed that listening comprehension is critical to reading comprehension (Foorman, Koon, Petscher, Mitchell, & Truckenmiller, 2015; Kim, 2015; Kim, Wagner, & Lopez, 2012; Kim & Wagner, 2015). When children present with reading comprehension deficiencies, there is a heavy focus on word recognition difficulties, including dyslexia and learning disabilities. Difficulties with word recognition are linked to weakness in the phonological domain of language and are often identified early on in the pre-school years (Catts, Fey, Zhang, & Tomblin, 2001). On the other hand, some children demonstrate reading comprehension difficulties despite adequate word reading abilities (Catts, Adlof, & Ellis Weismer, 2006; Nation, Clarke Marshall, & Durand, 2004). This group of individuals is known as poor comprehenders. Poor comprehenders are able to read text accurately and fluently at age-appropriate levels, however, they have difficulty understanding what they are reading (Cain & Oakhill, 2007; Nation, 2005). For example, when reading, poor comprehenders have weaknesses in the areas of semantics, syntax (Catts, Adlof, & Ellis Weismer, 2006; Nation & Snowling, 1998; Nation, Snowling, & Clarke, 2007) and more complex parts of language such as idioms, inferencing, comprehension monitoring, and knowledge of text structure (Oakhill, 1984; Cain & Towse, 2008; Cain, Oakhill, & Bryant, 2004; Oakhill & Yuill, 1996). Additionally, when we consider narrative comprehension, children with language disorders are less likely to provide correct answers to literal or inferential questions about stories that have been read to them (Gillam, Fargo, & Robertson, 2009; Laing & Kamhi, 2002). Since reading comprehension takes time to develop, it is difficult to demonstrate reading comprehension deficits in children before they are able to read accurately and fluently. Thus, these students' reading comprehension deficits may go unnoticed until later grades. As such, it is critical that language deficits are identified as early on in development as possible.

There is also a strong relationship between oral language abilities and reading ability (Hulme & Snowling, 2013). Nation, Clarke, Marshall, and Durand (2004) investigated poor compehenders' spoken language skills. The results of this study found that these students were less skilled than those in the typically developing group on semantic tasks (e.g., vocabulary and word knowledge), morphosyntax (e.g., past tense inflection, sentence comprehension) and aspects of language use (e.g., understanding figurative language). Research also suggests that students with expressive language difficulties are four

to fives times more likely than their peers to present with reading difficulties (Catts, Fey, Zhang, & Tomblin, 2001). For example, Zielinkski, Bench, and Madsen (1997) explored expressive language delays in preschoolers and found that these children were more likely to have difficulties with reading performance. Poll and Miller (2013) also reported that when children are 8 years old, expressive language delays could be a significant risk factor for poor oral language and reading comprehension. Furthermore, Lee (2011) discovered that expressive language development predicts comprehension of reading passages in both third and fifth grade students. Vocabulary can also play an important role early on in development as was demonstrated in Duff, Reen, Plunkett, and Nation's (2015) study that found infant vocabulary between 16 and 24 months is predictive of reading comprehension early on in school instruction years. Additionally, Pysridou, Eklund, Poikkeus, and Torppa's study (2018) found that expressive language ability at age 2–2.5 years old is associated with reading comprehension in ages 8–16 years old.

Listening comprehension and oral language abilities can also be important when we consider writing development (Kim, Al Otaiba, Wanzek, & Gatlin, 2015; Hulme & Snowling, 2013). Children with language impairments have been found to show grammatical errors (Gillam & Johnston, 1992; Scott & Windsor, 2000; Windsor, Scott, & Street, 2000) and spelling errors in their written texts. The spelling errors are similar to those found in children with dyslexia (Puranik, Lombardino, & Altmann, 2007), however, an individual's ability to create and think of new ideas appears to be specific to difficulties within the language system (Bishop & Clarkson, 2003; Puranik, Lombardino, & Altmann, 2007). Numerous studies have explored the difficulties that school-age children with language impairment have with telling stories. For example, when compared to typically developing children, children with language deficits tend to compose stories that contain fewer words and utterances (Moyano & McGillivray, 1988 [as cited in Hughes, McGillivray, & Schmidek, 1997]), fewer story grammar components (Paul, 1996), reduced sentence complexity (Gillam & Johnston, 1992), fewer complete cohesive ties (Liles, 1985), increased grammatical errors (Liles, Duffy, Merritt, & Purcell, 1995; Norbury & Bishop, 2003), and poorer overall story quality (Gillam, McFadden, & van Kleeck, 1995; McFadden & Gillam, 1996).

Over the last thirty years, there has been an abundance of research demonstrating that children with specific language impairment (SLI) are at a disadvantage when it comes to peer relationships (Durkin & Conti-Ramsden, 2010). Individuals with SLI engage less in active conversation interactions, enter less frequently into positive social interactions, demonstrate poorer discourse skills, are more likely to provide inappropriate verbal responses, and are less likely to influence others successfully (Hadley and Rice, 1991; Craig, 1993; Craig and Washington, 1993; Grove, Conti-Ramsden, & Donlan, 1993; Guralnick, Connor, Hammond, Gottman, & Kinnish, 1996; Brinton, Fujiki, & McKee 1998; Vallance, Im, & Cohen 1999). Children with SLI also tend to score lower in the areas of social skills, social cognitive abilities, and may have trouble with behavioral and emotion regulation (Cohen, Barwick, Horodezky, Vallance, & Im, 1998; Fujiki, Brinton, & Clarke, 2002; Marton, Abramoff, &Rosenzweig, 2005; Lindsay, Dockrell, & Strand, 2007). Additionally, children with language impairments are at higher risk of academic failure, social exclusions, behavioral and emotional difficulties, and are more vulnerable to being bullied (Conti-Ramsden, Durkin, Simkin, & Knox, 2009; St Clair, Pickles, Durkin, & Conti-Ramsden, 2011). Lastly, children with language disorders are also at a heightened risk of exhibiting externalizing problems and antisocial conduct disorders (Beitchman, Wilson, Johnson, et al., 2001; Conti- Ramsden & Botting, 2004).

Restating Information Test

Why is restating information important to assess?

When children are asked to listen to and then retell a story in their own words, they are providing an oral narrative by retelling past events from their own perspective using chronological and causal sequences of events (Gillam & Pearson, 2004; Burdelski & Evaldsson, 2019; Takagi, 2019). This story retell task provides clinicians with narrative samples that are important to include in a comprehensive speech and language assessment because these skills are closely related to later reading comprehension (Griffin, Hemphill, Camp, & Wolf, 2004), academic achievement (Fazio, Naremore, & Connel, 1996), and performance on norm-referenced assessments of language (Ebert & Scott, 2014; Scott & Windsor, 2000).

Children's ability to recall the sequence of events relies largely on the way the events were originally encoded. When children are presented with a causal sequence, they are more likely to remember an event (van den Broek, 1997). The encoding process develops and grows over time. For example, when children are around 4 years of age, they can label salient details, at 5 years of age, they begin to include some temporal sequence of events, and at 6 years old children begin to discuss causal relationships (Bishop & Donlan, 2005). Later, around 10 years old or older, children begin to include psychological causality and characters' motivations. Research has suggested that when causality information is encoded, the story is more resistant to forgetting than when compared to a fragmented series of details (Bishop & Donlan, 2005).

When documenting a child's oral narratives, clinicians have the opportunity to examine vocabulary, grammar, and narrative organizational skills (Westby, 2005). Microstructural analyses can take a look at children's grammatical and syntactic abilities and can be analyzed through mean length of utterance (Miller, 1981) and sentence complexity (Nippold, Hesketh, Duthie, & Mansfield, 2005; Schuele & Tolbert, 2001). Children's vocabulary can also be analyzed by calculating lexical diversity and number of different words (Miller & Klee, 1995). Macrostructural analyses can investigate how children relate concepts. These analyses assume the setting and episode system where the setting provides background information about the characters and their environments and the episode systems includes a problem, solution, and conclusion (Mandler & Johnson, 1977; Rumelhart, 1975; Stein & Glenn, 1979).

Children with language impairment may have difficulty using appropriate vocabulary and grammar when telling stories and difficulty with text-level organization of narratives (Pearce, McCormack, & James, 2003; Reilly, Losh, Bellugi, & Wulfeck, 2004). For example, if a child has difficulty with vocabulary and grammar, it will be difficult for him/her to produce fully competent utterances and if the child has impaired macrostructural skills, he/she will have difficulty generating coherent and age-appropriate extended discourse (Heilmann, Miller, Nockerts, & Dunaway, 2010).

Following Directions Test

Why is following directions important to assess?

In the classroom, students are constantly receiving information and being asked to follow instructions (both auditorily and visually) from their teachers. A child's ability to understand and follow verbal directions is a fundamental developmental skill that is necessary to learn and succeed in the classroom (Gill, Moorer-Cook, Armstrong, & Gill, 2012). Following verbal instructions requires many interrelated skills such as the ability to hear the instruction, understand the vocabulary and comprehend the syntax,

and utilize working memory (MacDonald & Christiansen, 2002; Alloway, Gathercole, Willis, & Adams, 2004). The ability to comprehend directions is a skill that is often used to determine receptive language difficulties (Catts, Fey, Tomblin, & Zhang, 2002; Mainela, Evans, & Coady, 2008; Seigner-Gardner & Schwartz, 2008), whereas the ability to follow spoken directions is a skill that could be used to indicate auditory processing or receptive language difficulties (Ferguson, Hall, Riley, & Moore, 2011). In order to provide appropriate assessment and intervention to our students, is important to understand the cognitive processes and language factors that are required for these verbal tasks (Archibald, 2013).

Research has demonstrated that a child's ability to understand sentences of increasing length does not always relate to syntactic maturity (Nippold, 2007). For example, many structures may increase the complexity of language (e.g., subordinate clauses, participle phrases) without actually increasing the length of the sentence. Robertson and Joanisse's (2010) study demonstrated that sentence length as well as syntactic complexity may impact a child's ability to comprehend a spoken sentence. Using a picture pointing task, Robertson and Joanisse (2010) found that typically developing children, children with dyslexia, and children with language impairment had more difficulty processing longer versus shorter sentences. Additionally, the study showed that children with language impairment had more difficulty processing sentences with passive versus active voice. Lastly, the researchers found an interaction effect across all groups where children had more difficulty processing syntactically complex sentences that were in longer versus shorter sentences. Thus, it can be presumed that children with language impairments may have greater difficulty following directions as they become increasingly longer or more syntactically complex.

When we consider working memory, sometimes, the informational load a child is given can be too much for what his/her working memory can handle at that given moment. This can lead to information loss because the student cannot hold that much information in their working memory. When this happens, working memory performance can negatively impact a student's ability to follow directions (Bergman-Nutley & Klingberg, 2014). Allen and Waterman (2015), suggest that in order to assist students, we can ask students to immediately act on the received information. The ability to recall instructions increases when the student is asked to do so immediately after instruction. A second strategy to assist students is to use different forms of instructions (e.g., written, verbal, visual) (Mayer, 2008).

Morphology and Sentence Structure Test

Why is morphology and sentence structure important to assess?

Beginning as early as preschool and continuing into the school-age years, children are required to listen to information that frequently uses complex syntax. Additionally, children are expected to use complex syntax when they speak. Complex syntax involves utterances that contain more than one clause, through coordination or subordination (Quirk, Greenbaum, Leech, & Svartik, 1985). For typical language learners, complex syntax production begins shortly after children begin to combine words and continues to develop rapidly across the preschool years into the school-age years (Barako Arndt & Schuele, 2013). Preschool-age children begin with acquiring grammatical morphemes and other basic clausal structures (simple sentences). Research has demonstrated that children with speech or language impairment (SLI) produce fewer instances of complex syntax (Marinellie, 2004) and may leave out required grammatical elements (e.g., subject relative markers, nonfinite to markers) (Barako Arndt & Schuele, 2012; Owen & Leonard, 2006; Schuele & Dykes, 2005).

As expected, a child's ability to understand and use complex syntax will influence academic success (Scott & Windsor, 2000). For example, complex syntax is required to engage in classroom discourse,

answer questions, summarize and explain information, and also used for written expression and used socially in conversations. For some early talkers, complex syntax may emerge by 2 years of age (Diessel, 2004). Typically, children are regularly using complex syntax between their second and third birthdays and are capable of using a variety of complex syntax types by the time they begin kindergarten (Bloom, Tackeff, & Lahey, 1984). Complex syntax production continues to expand and grow throughout the school-age years in order to meet classroom demands and communicative expectations. If syntax production is not appropriately assessed and treated early on, children may not be able to meet the language comprehension and production demands found in kindergarten or first grade (Barako Arndt & Schuele, 2013).

Researchers have suggested that conversational samples can be used successfully to assess children's language production abilities. For example, play-based strategies can be used with preschoolers (e.g., play house, farm), topics of interest (e.g., movies, sports, school activities) can be used with early elementary school-age children, and a variety of discourse tasks (e.g., conversation, narrative, expository) can be used with older children (Hadley, 1998; Barako Arndt & Schuele, 2013). Children can be asked to describe or to explain how to complete a task such as a cooking recipe or playing a game (Nippold, 2010) or to summarize a video (Scott & Windsor, 2000).

Chapter

3

Administration and Scoring Procedures

The following testing guidelines represent specific administration and scoring procedures for the *Language Video Assessment Tool.* These procedures are considered best professional practice required in any type of assessment as described in the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, and NCME], 2014).

Examiner Qualifications

Professionals who are formally trained in the ethical administration, scoring, and interpretation of assessment tools and who hold appropriate educational and professional credentials may administer the *Language Video Assessment Tool.* Qualified examiners include speech-language pathologists and clinical fellows in speech language pathology. It is a requirement to read and become familiar with the administration, recording, and scoring procedures before using this assessment.

Confidentiality Requirements

As described in Standard 6.7 of the Standards for Educational and Psychological Testing (AERA et al., 2014), it is the examiner's responsibility to protect the security of all testing material and ensure confidentiality of all testing results.

Target Population for Testing

The standardization process undertaken by the *Language Video Assessment Tool* allows it to be used for individuals between the ages of 5-18. The *Language Video Assessment Tool* provides information regarding an individual's spoken language comprehension, oral expression and language integration. Students with these difficulties will be brought to the attention of speech-language pathologists, school psychologists, teachers, parents, reading specialists and others who are involved with the academic and social impact of language impairments.

The Language Video Assessment Tool can be used to identify language impairment, help aid in the eligibility criteria needed for special education services or can be used to provide a description of current language skills. The target populations for this assessment are provided below:

- ➔ Students who have been previously diagnosed or are suspected of having a spoken language

comprehension or expressive language disorder (also known as specific language impairment, developmental language disorder, speech or language impairment, or language learning disability). This assessment can help provide criteria for a language impairment diagnosis and/or eligibility.

- ➔ Students suspected of having a learning disability in the area of oral or written language (also known as specific learning disability, language-based learning disability, or language learning disability). This assessment can help provide criteria for language impairment diagnosis and/or eligibility.
- ➔ Students with known diagnoses, such as intellectual disability, autistic spectrum disorder, intellectual disorder, and traumatic brain disorder. This assessment can help provide a description of current language skills.
- ➔ Students with known difficulty of language, literacy, or social communication skills that have not met criteria for a formal diagnosis but are receiving support services. This assessment can help provide a description of current language skills.
- ➔ Students in any of the above groups, who have been previously assessed with *The Language Video Assessment Tool,* can also benefit with a follow-up test for the purpose of tracking progress. Note: Follow-up assessments should take place at least 6-months or later after the previous assessment date.

Testing Time

Administration of the test takes approximately 40-50 minutes.

Additional Testing Considerations and Procedures

A. Seating arrangement is important when administering this test because both the examiner and the student need to be able to see the videos. The examiner must be able to face the student during testing in order to closely observe his/her use of articulation.

B. Administer the test in a quiet, comfortable environment with no distractions. Stop testing if the student appears to be tired or is unwilling to participate.

C. It is important to elicit the examinees' <u>best effort</u> on each test and on each item presented. This can be achieved by establishing rapport with the examinee before the testing begins and by providing praising prompts when needed.

D. Because this is not a timed test, examinees should be allowed time to respond. However, if no response is provided within 10 seconds of presentation of an item, the clinician should prompt the examinee to imitate the target word.

E. If the examiner has reasons to believe that the testing results are invalid, such as poor attention span that is noticeably different from those expected, or student showing sign of being ill, retest at a later time.

Morphology

Repetition of Video Test Items

Repetition of videos is allowed up to two times.

Repetition of Item Questions

Repetition of item questions is allowed when the examinee appears to not understand the question or requests a repetition. If the examinee does not provide a response after the second reading, score the test item as a "0" and proceed to the next test item.

Prompting Rules

The Morphology and Sentence Structure test items were designed to be easily recognizable by young children and older students. However, when responses are ambiguous or other than the target word, the examiner should attempt to elicit the target response by repeating the test item question. If the examinee does not provide a response after the second reading, score the test item as a "0" and proceed to the next test item.

Language Comprehension

Repetition of Video Test Items

Repetition of videos is not allowed unless it is apparent that the background noise interferes with comprehension of video content.

Repetition of Item Questions

One repetition of each item question is allowed when the examinee appears to not understand the question or requests a repetition. If the examinee does not provide a response after the second reading, proceed to the next question.

Prompting Rules

If the initial response to a test item is ambiguous, the examiner should repeat the question. If the examinee does not provide a response after the repetition, the examiner should proceed to the next test item.

Restating information

Repetition of Video Test Items

Repetition of videos is not allowed unless it is apparent that the background noise interferes with comprehension of video content.

Prompting Rules

If the examinee does not provide a sufficient response or does not offer a sufficient amount of information and details, the examiner should prompt by saying, “Tell me more” or “Tell me anything you remember from the video.”

Following Directions

Repetition of Item Questions/ Prompting Rules

Repetition of item questions is allowed when the examinee appears to not understand the question or requests a repetition. If the examinee does not follow the directive after the second reading, score the test item as a “0” and proceed to the next test item.

Test Materials

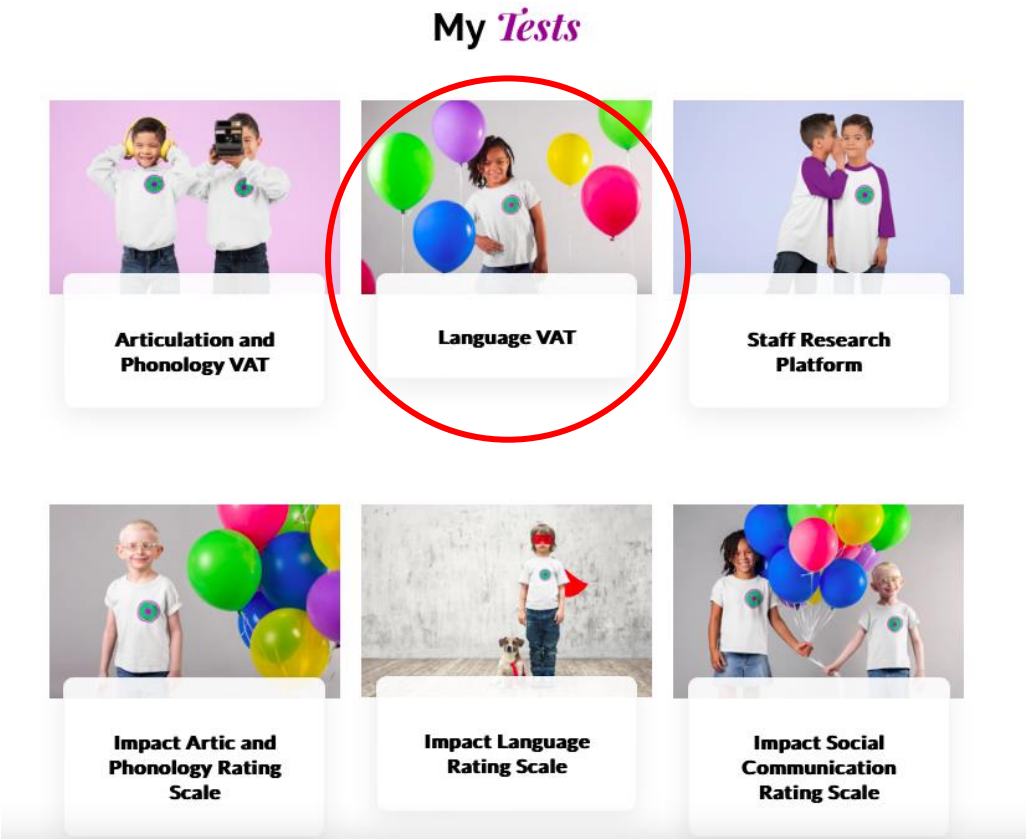
The *Language Video Assessment Tool* consists of four independent tests: Language Comprehension, Restating Information, Following Directions and Morphology & Sentence Structure. All video-based test items, paper-based or digital protocols, and scale converting software is available on the *Video Assessment Tools* website at: www.videoassessmenttools.com

Administration Instructions

Begin by logging onto your Video Assessment Tools account: www.videoassessmenttools.com

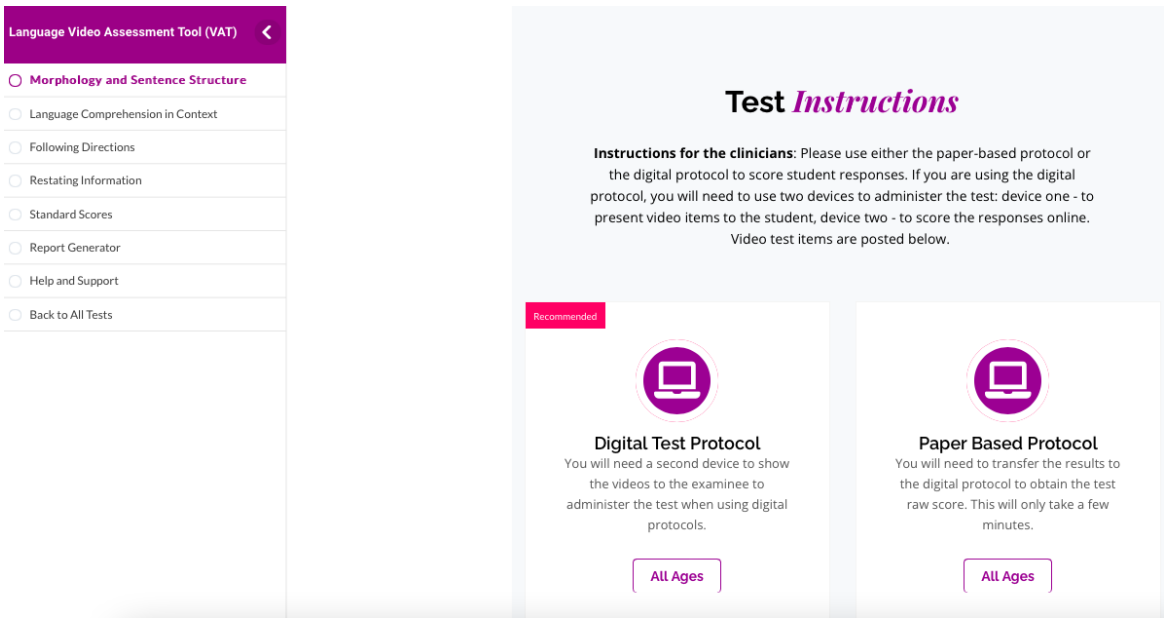
Next, select the “Administer Tests” tab and scroll down to the *Articulation and Phonology VAT*.

Select the *Language VAT* by clicking on the picture.



Administration Instructions

Step 1: After you have selected the *Language VAT*, you will see a toolbar on the left of your screen. Select the name of the subtest you would like to administer first. Here, you will also see a "Digital Test Protocol" and "Paper Based Protocol" tab. Clinicians may choose to use the digital based protocol or the paper-based protocol when scoring. This choice is made based on the comfort level and/or preference of the examiner.



If clinicians do choose the hardcopy paper protocol, the results will have to be transferred to the digital protocol after completing the assessment online to obtain the raw scores automatically. Transferring data from the hardcopy protocol to the digital protocol is a quick and easy process and should take no more than five minutes.
*When accessing the protocol, you will see a “Download protocol here” tab. Select this tab to download a PDF copy which can be worked on online or printed.

Step 2: Once the assessment loads, there will be two viewing options available. *Option A* allows you to use buttons to navigate through videos. *Option B* allows you to view videos individually. Both options contain the same videos and test items, choose the option you prefer. For example, you may prefer *Option B* if the test is being administered over the course of a few days.

Step 3: Instructions for each test:

Language Comprehension in Context
This test does not require a trial test item. Instead, the examinee is asked to repeat what they are being asked to do. Please ensure that the student understands that they are being asked to watch videos and answer questions based on the videos. At the beginning of the test, the clinician explains the test to the examinee by saying, *"You are about to watch some videos about nature. Please watch and listen carefully. You will be allowed to watch each video only one time. I will ask you questions about each video. Please do your best to answer the questions. You can ask me to repeat the question. I will be able to repeat the question only one time. If you don't know or can't remember the answer, it's okay to say, ‘I don't know’ or ‘I can't remember.’ So please tell me now, what are you supposed to do?"*
Then the clinician will play the video, and the examinee will be asked questions about what happened/what was discussed in the video clip.

Ceiling Rule for the “BEES" video section: 6 consecutive “0” scores. Proceed to the "SEALIFE" videos after 6 consecutive “0” scores.
Ceiling Rule for the “SEALIFE" video section: Discontinue administering the test after 6 consecutive “0” scores. Please note that this form does not allow you to "unclick" choices.

Following Directions Test
At the beginning of the test, the clinician explains the test to the examinee by saying, *"I am going to show you some pictures and ask you to draw some things. First, I want you to listen while I ask you to do something. Then, after I say ‘go,’ you can follow the instructions and draw. I can repeat the instructions for you one time if you need me to. So, can you please tell me what you are supposed to do now?"*

Scoring: Clinicians must score each test item as “1” when it is executed correctly. All test directives executed out of order or incorrectly must be scored as a “0”.
Ceiling Rule for "Executing Directions" section: 5 consecutive “0” scores. Proceed to the "Following Verbal Directions" part after 5 consecutive “0” scores.
Ceiling Rule for "Verbal Directions" section: Discontinue administering the test after 5 consecutive “0” scores. Please note that this form does not allow you to "unclick" choices.

Restating Information

At the beginning of the test, the clinician explains the test to the examinee by saying, ***"You are going to watch some videos. Please watch and listen carefully. You will watch the video only one time. After you are done watching, I need you to tell me what the videos are about and all the details that you can remember. So, tell me, what are you supposed to do now?"*** (Please ensure that the examinee understands that they are being asked to watch videos and answer questions based on the video). Proceed to the first video. Once the video finishes playing, say, "Now tell what the video was about and all the details you can remember".

Prompting: If the student does not provide a sufficient response or does not offer a sufficient amount of information and details, say, "Tell me more" or "Tell me anything you remember from the video."
Basal/Ceiling Rule: There is no basal/ceiling rule on this test. Please note that this form does not allow you to "unclick" choices.

Morphology and Sentence Structure

At the beginning of the test, the clinician explains the test to the examinee by saying, *"You are going to watch some videos. After watching each video, you will be asked to answer a question about the video or you will be asked to complete a sentence about the video. Let's try one."* Please play the trial video test item. Allow the student to answer. If the student doesn't answer correctly, say*, "He is reading."* Proceed with the test items. If needed, you may repeat showing the videos up to two times. Correct test item answers are provided for clinicians' reference.

Basal/Ceiling Rule: None. Please administer all test items. Please note that this form does not allow you to "unclick" choices.

Step 5: Next, select the "Get Standard Scores" from the side toolbar. Here you will enter the student's raw score and age to arrive at a standard score and percentile rank.

Step 6: Lastly, use the optional report generator to assist you in writing the report for the *Language VAT*.

Chapter

4

Standardization and Normative Information

The normative data for the *Language Video Assessment Tool* is based on the performance of 1554 examinees across 11 age groups (shown in Table 4.1) from 17 states across the United States of America (Arizona, California, Colorado, Nevada, Idaho, Illinois, Iowa, Kansas, Ohio, Minnesota, Florida, New York, Pennsylvania, Florida, South Carolina, Texas, Washington).

Table 4.1
Representation of the Sample, by Age Group

Age Group	Age	*N*	%
1	5-0 to 5-11	163	8.4
2	6-0 to 6-11	142	8.7
3	7-0 to 7-11	151	9.4
4	8-0 to 8-11	139	9.9
5	9-0 to 9-11	134	9.1
6	10-0 to 10-11	129	8.5
7	11-0 to 11-11	137	8.2
8	12-0 to 12-11	119	8.4
9	13-0 to 13-11	141	9.0
10	14-0 to 14-11	136	9.4
11	15-0 to 18-0	163	11.0
Total		1554	100%

The data was collected throughout the 2016-2020 school years by 34 state licensed speech-language pathologists (SLPs). The SLPs were recruited through Go2Consult Speech and Language Services, a speech-language pathology services and nonpublic agency certified by the CA Department of Education in conjunction with the Lavi Institute, an ASHA approved CE provider. All standardization project procedures were reviewed and approved by IntegReview IRB (now known as Advarra), a fully AAHRPP-accredited independent review board that provides ethical review for all phases of industry-sponsored and federally funded research in the U.S. To ensure representation of the national population, the *Language Video Assessment Tool* standardization sample was selected to match the US Census data

reported in the ProQuest Statistical Abstract of the United States (ProQuest, 2017). The sample was stratified within each age group by the following criteria: gender, race or ethnic group, and geographic region. The demographic table below (Table 4.2) specifies the distributions of these characteristics and shows that the normative sample is nationally representative.

Table 4.2

Demographics of the Normative Sample vs. US Population

Normative Sample Size = 1554

Demographic	*N* Normative Sample	% Normative Sample	% US Population
Gender			
Male	770	49.5%	49%
Female	784	50.5%	51%
Total	1554	100%	100%
Race			
White	963	62%	77%
Black	233	15%	13%
Asian	78	5%	4%
Hispanic	202	13%	12%
Other	78	5%	6%
Total	1554	100%	100%
Clinical Groups			
	none	none	none
US Regions			
Northeast	279	18%	16%
Midwest	295	19%	22%
South	528	34%	38%
West	452	29%	24%
Total	1554	100%	100%
Parents' Educational Level			
Four years of college or more	466	30	31%
Some college	404	26	27%
High school graduate	482	31	30%
Less than high school graduate	202	13	12%
Total	1554	100%	100%

Criteria for inclusion in the normative sample

A strong assessment is one that provides results that will benefit the individual being tested or society as a whole (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, and NCME], 2014). One way we can tell if an assessment is strong, is if the test includes adequate norms. Previous research has suggested that utilizing a normative sample can aid in the identification of a disability. Research has also suggested that the inclusion of children with disabilities may negatively impact the test's ability to differentiate between children with disorders and children who are typically developing (Peña, Spaulding, & Plante, 2006). Since the purpose of the *Language Video Assessment Tool* is to help to identify students who present with language disorders, it was critical to exclude students from the normative sample who have diagnoses that are known to influence language abilities (Peña, Spaulding, & Plante, 2006). Students who had previously been diagnosed with spoken language comprehension and/or spoken language disorders, auditory processing disorders, and articulation or phonological impairments were not included in the normative sample. Further, students were excluded from the normative sample if they were diagnosed with autism spectrum disorder, intellectual disability, hearing loss, neurological disorders, or genetic syndromes. In order for students to be included in the normative sample for this assessment tool, students must have met criteria of having typical language development, and show no evidence of language deficits. Thus, the normative sample for the *Language Video Assessment Tool* provides an appropriate comparison group (i.e., a group without any known disorders that might affect language abilities) against which to compare students with suspected disorders.

The *Language Video Assessment Tool* is designed for students who are native speakers of English and/or are English language learners (ELL) who have demonstrated a proficiency in English based on state testing scores and school district language evaluations. Additionally, students who were native English speakers and also spoke a second language were included in this sample.

Norm-referenced testing is a method of evaluation where an individual's scores on a specific test are compared to scores of a group of test-takers (e.g., age norms) (AERA, APA, and NCME, 2014). Clinicians can compare clinician, teacher, and parent ratings on the *Language Video Assessment Tool* to this normative sample to determine whether a student is scoring within normal limits or, if their scores are indicative of a language disorder. Administration, scoring, and interpretation of the *Language Video Assessment Tool* must be followed in order to make comparisons to normative data. This manual provides instructions to guide examiners in the administration, scoring, and interpretation of the assessment results.

Chapter

5

Validity and Reliability

This section of the *Language Video Assessment Tool* manual provides information on the psychometric characteristics of validity and reliability. Validity helps establish how well a test measures what it is supposed to measure and reliability represents the consistency with which an assessment tool measures a certain ability or skill. The first half of this chapter will evaluate content, construct, criterion, and clinical validity of the *Language Video Assessment Tool*. The latter half of the chapter will review the consistency and stability of the *Language Video Assessment Tool* scores, in addition to test retest and inter-rater reliability.

Validity

Validity is essential when considering the strength of a test. Content validity refers to whether the test provides the clinician with accurate information on the ability being tested. Specifically, content validity measures whether or not the test actually assesses what it's suppose to. According to McCauley and Strand (2008), there should be a rationalization of the methods used to choose content, expert evaluation of the test's content, and an item analysis.

Content-oriented evidence of validation addresses the relationship between a student's learning standards and the test content. Specifically, content-sampling issues look at whether cognitive demands of a test are reflective of the student's learning standard level. Additionally, content sampling may address whether the test avoids inclusion of features irrelevant to what the test item is intended to target.

Single-cut Scores

It is common to use single cut scores (e.g., -1.5 standard deviations) to identify disorders, however, there is evidence that advises against using this practice (Spaulding, Plante, & Farinella, 2006). When using single cut scores (e.g., -1.5 SD, -2.5 SD, etc.), we may under identify students with impairments on tests for which the best-cut score is higher and over identify students' impairments on tests for which the best-cut score is lower. Additionally, using single cut scores may go against IDEA's (2004) mandate, which states assessments must be valid for the purpose for which they are used.

Inclusion/Exclusion Criteria for the Discriminant Analysis and the Group Differences Study

Typically developing participants were selected based on the following criteria: 1) exhibited hearing sensitivity within normal limits; 2) presented with age-appropriate speech and language skills; 3) successfully completed each school year with no academic failures; and 4) attended public school and placed in general education classrooms.

Inclusion criteria for the developmental language disorder - language comprehension group was: 1) having a current diagnosis of developmental language disorder - language comprehension impairment (based on medical records and/or school-based special education eligibility criteria); 2) being enrolled in the general education classroom for at least 4 hours per day; and 3) exhibited hearing sensitivity within normal limits.

Finally, the inclusion criteria for the developmental language disorder - expressive language impairment group was: 1) having a current diagnosis of a developmental language disorder - expressive language impairment or delay (based on medical records and/or school-based special education eligibility criteria); 2) being enrolled in the general education classroom for at least 4 hours per day; and 3) exhibited hearing sensitivity within normal limits.

Sensitivity and Specificity

Table 5.1 shows the cut scores needed to identify language disorders within each age range. Additionally, this table demonstrates the sensitivity and specificity information that indicates the accuracy of identification at these cut scores. Sensitivity and specificity are diagnostic validity statistics that explain how well a test performs. Vance and Plante (1994) set forth the standard that for a language assessment to be considered clinically beneficial, it should reach at least 80% sensitivity and specificity. Thus, strong sensitivity and specificity (i.e., 80% or stronger) is needed to support the use of a test in its identification of the presence of a disorder or impairment. Sensitivity measures how well the assessment will accurately identify those who truly have a language disorder (Dollaghan, 2007). If sensitivity is high, this indicates that the test is highly likely to identify the language disorder, or, there is a low chance of "false positives." Specificity measures the degree to which the assessment will accurately identify those who do not have a language disorder, or how well the test will identify those who are "typically developing" (Dollaghan, 2007).

Table 5.1 Language Video Assessment Tool sensitivity, specificity and likelihood ratios

Language Comprehension Test

Age group	Cut score	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio
5:0-5:11	77	.87	.88	4.04	.09
6:0-6:11	77	.81	.94	5.05	.11
7:0-7:11	77	.91	.83	4.02	.23
8:0-8:11	78	.88	.81	4.12	.41
9:0-9:11	77	.84	.92	4.08	.23
10:0-10:11	77	.92	.91	5.42	.16
11:0-11:11	78	.91	.89	6.07	.06
12:0-12:11	77	.88	.92	4.06	.23
13:0-13:11	77	.89	.83	5.71	.18
14:0-14:11	78	.93	.90	4.03	.26
15:0-15:11	78	.89	.91	6.11	.08
16:0-21:0	78	.94	.96	4.23	.21

Restating Information Test

Age group	Cut score	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio
5:0-5:11	77	.88	.91	4.04	.09
6:0-6:11	77	.89	.87	4.03	.14
7:0-7:11	77	.92	.86	4.34	.13
8:0-8:11	77	.89	.93	5.04	.12
9:0-9:11	77	.92	.91	4.84	.15
10:0-10:11	77	.86	.84	4.04	.18
11:0-11:11	78	.81	.87	4.01	.32
12:0-12:11	77	.92	.91	5.67	.08
13:0-13:11	77	.88	.86	4.04	.34
14:0-14:11	78	.84	.92	4.23	.23
15:0-15:11	77	.84	.88	6.72	.13
16:0-21:0	77	.82	.92	5.12	.15

Note: Age groups 16:0-21:0 are reported together as there were no age-related changes detected after the age of 16. Total N=3037; typically developing group n=1554; clinical group=1483

Table 5.1 Language Video Assessment Tool sensitivity, specificity and likelihood ratios

Following Directions Test

Age group	Cut score	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio
5:0-5:11	77	.88	.91	4.03	.08
6:0-6:11	77	.91	.90	4.11	.11
7:0-7:11	78	.88	.84	4.23	.21
8:0-8:11	77	.86	.83	4.21	.09
9:0-9:11	78	.94	.85	5.84	.11
10:0-10:11	77	.87	.93	4.89	.23
11:0-11:11	78	.85	.91	4.21	.31
12:0-12:11	78	.90	.92	4.11	.46
13:0-13:11	77	.93	.83	4.16	.08
14:0-14:11	77	.89	.91	4.14	.12
15:0-15:11	77	.87	.83	5.23	.24
16:0-21:0	78	.85	.82	4.56	.40

Morphology and Sentence Structure Test

Age group	Cut score	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio
5:0-5:11	77	.88	.84	4.01	.09
6:0-6:11	78	.83	.91	4.10	.11
7:0-7:11	77	.87	.83	4.24	.34
8:0-8:11	77	.91	.80	5.09	.06
9:0-9:11	78	.93	.85	6.34	.23
10:0-10:11	77	.88	.89	5.07	.12
11:0-11:11	77	.83	.82	4.67	.08
12:0-12:11	78	.92	.85	4.12	.12
13:0-13:11	77	.89	.81	4.23	.11
14:0-14:11	77	.93	.89	4.07	.09
15:0-15:11	77	.91	.90	5.89	.08
16:0-21:0	77	.88	.84	6.11	.44

Note: Age groups 16:0-21:0 are reported together as there were no age-related changes detected after the age of 16. Total N=3037; typically developing group n=1554; clinical group=1483

Content Validity

The validity of a test determines how well the test measures what it purports to measure. Validity can take various forms, both theoretical and empirical. This can often compare the instrument with other measures or criteria, which are known to be valid (Zumbo, 2014). For the content validity of the test, expert opinion was solicited. Thirty-one speech language pathologists (SLPs) reviewed the *Language Video Assessment Tool*. All SLPs were licensed in the state of California, held the Clinical Certificate of Competence from the American Speech-Language-Hearing Association, and had at least 5 years of experience in assessment of children with spoken language comprehension, spoken, and social language disorders. Each of these experts was presented with a comprehensive overview of each test descriptions, as well as rules for standardized administration and scoring. They all reviewed 6 full-length administrations. Following this, they were asked 35 questions related to the content of the tests and whether they believed the assessment tool to be an adequate measure of language functioning. For instance, their opinion was solicited regarding whether the questions and the examinees' responses properly evaluated language comprehension and oral expression skills. The reviewers rated each test on a decimal scale. All reviewers agreed that the *Language Video Assessment Tool* is a valid standardized measure to evaluate language skills in students who are between the ages of 5 and 18 years old. The mean ratings for the Language Comprehension, Restating Information, Following Directions and Morphology & Sentence Structure tests were 31.4±0.6, 29.3±0.4, 29.4±0.9, 30.3±0.7, respectively.

Construct Validity

Developmental Progression of Scores

Spoken language comprehension and spoken language is developmental in nature and skills change with age. Mean raw scores for examinees should increase with chronological age, demonstrating age differentiation. Mean raw scores and standard deviations for the *Language Video Assessment Tool* are divided into eleven age intervals displayed in Table 5.2 below.

Table 5.2 Normative Sample's mean raw scores and standard deviations on the *Language Video Assessment Tool*

	Tests			
Age Group	Language Comprehension	Following Directions	Restating Information	Morphology and Sentence Structure
5:0-5:11	78 (4.1)	44 (3.1)	42 (4.1)	55 (2.6)
6:0-6:11	84 (2.7)	47 (2.9)	45 (3.4)	57 (3.1)
7:0-7:11	89 (2.5)	49 (3.4)	48 (3.9)	57 (3.4)
8:0-8:11	92 (2.8)	50 (2.8)	51 (2.6)	58 (2.9)
9:0-9:11	94 (3.1)	53 (2.6)	53 (2.4)	61 (2.6)
10:0-10:11	95 (3.2)	56 (3.1)	55 (2.1)	62 (1.9)
11:0-11:11	98 (2.1)	56 (2.7)	57 (2.4)	64 (1.7)
12:0-12:11	98 (1.9)	57 (1.8)	57 (1.8)	64 (2.1)
13:0-13:11	99 (1.4)	57 (1.4)	58 (1.4)	65 (1.8)
14:0-14:11	100 (1.1)	57 (0.9)	60 (1.7)	65 (1.6)
15:0-15:11	100 (0.5)	58 (0.6)	60 (1.5)	66 (1.4)
16:0-21:0	100 (0.3)	58 (0.7)	63 (1.7)	66 (1.7)

Group Differences

Since a language assessment tool is designed to identify those examinees with developmental language disorder, it would be expected that individuals identified as likely to exhibit language comprehension and/or oral expression difficulties would score lower than those who are typically developing. The mean for the outcome variables (Language Comprehension, Restating Information, Following Directions and Morphology & Sentence Structure) were compared among the two clinical groups and the typically developing group of examinees using Kruskal Wallis analysis of variance (ANOVA). The level of significance was set at $p \leq 0.05$. Table 5.3 reviews the ANOVA, which reveals a significant difference between all three groups.

Table 5.3: Raw Score Means (and Standard Deviations) of each test for Two Clinical Groups and a Demographically Matched Typically Developing Group, (N=263)

	DLD-C group (n=89)	DLD-E group (n=83)	TD group (n=91)	p-value*
Language Comprehension [a,b,c]	69(4.2)	73(5.7)	98(3.2)	<.001
Restating Information [a,b,c]	47(5.6)	45(3.8)	57(2.6)	<.001
Following Directions [a,b,c]	48(4.7)	47(3.6)	56(2.4)	<.001
Morphology and Sentence Structure	48(3.8)	47(4.3)	64(2.7)	<.001

Abbreviation: DLD-C, Developmental Language Disorder - Language Comprehension; DLD-E, Developmental Language Disorder - Expressive Language; and TD, Typically Developing
*Kruskal-Wallis Analysis of Variance test
[a] significant difference between DLD-C and TD groups
[b] significant difference between DLD-E and TD groups
[c] significant difference between DLD-C and DLD-E groups

Standards for fairness

Standards of fairness are crucial to the validity and comparability of the interpretation of test scores (AERA, APA, and NCME, 2014). The identification and removal of construct-irrelevant barriers maximizes each test- taker's performance, allowing for skills to be compared to the normative sample for a valid interpretation. Test constructs and individuals or subgroups of those who the test is intended for must be clearly defined. In doing so, the test will be free of construct-irrelevant barriers as much as possible for the individuals and/or subgroups the test is intended for. It is also important that simple and clear instructions are provided.

Criterion Validity

In assessing criterion validity, a full correlation analysis was not possible for the Language VAT when compared to the current body of language tests. The Language Video Assessment Tool is unique in its content and design. Therefore, criterion validity should be analyzed with caution. The Language VAT cannot be expected to fully correlate to the existing body of language assessments because of its unique focus on functional language stimuli to resemble real-life academic presentations which is not available within other language tests.

To examine criterion validity, correlations of Language Comprehension, Restating Information, Morphology and Sentence Structure and Following Directions tests with two other measures of language were conducted. The Clinical Assessment of Spoken Language (*CASL)* is an individually-administered oral language assessment for students with ages 3 to 21 years. The test measures lexical, semantic, syntactic, and pragmatic language categories (Carrow-Woolfolk, 2017). The TILLS is a comprehensive, norm-referenced test that has been standardized for three purposes: To identify language/literacy disorders, to document patterns of relative strengths and weaknesses, and to track changes in language and literacy skills over time in students ages 6–18 years (Nelson et. al, 2016).

The Language Comprehension Test, the Meaning from Context subtest of the CASL test as well as the Language Comprehension subtest of the TILLS test were administered to 30 participants in counterbalanced order. Time between test administrations ranged from the same day to 5 days. The Restating Information Test, the Sentence Expression subtest of the CASL test and the Story Retelling subtest of the TILLS test were administered to 30 participants in a counterbalanced order. Time between test administrations ranged from the same day to 5 days. The Following Directions test and the Following Directions subtest of the TILLS test were administered to 30 participants in a counterbalanced order. Time between test administrations ranged from the same day to 5 days. The Morphology and Sentence Structure test and the Grammatical Morphemes subtest of the CASL test were administered to 30 participants in a counterbalanced order. Time between test administrations ranged from the same day to 5 days.

The concurrent validity was assessed using Pearson's correlation among the Language VAT, CASL and TILLS tests. Correlation coefficients of ≥ 0.7 are recommended for same-construct instruments while moderate correlations of ≥ 0.4 to ≤ 0.70 are acceptable. The level of significance was set at $p \leq 0.05$. When assessing validity, the Language VAT tests were substantially correlated with the Story Retelling, Language Comprehension and Following Directions subtests of the TILLs test. The correlation between the Language VAT tests and the CASL subtests were 0.69 and 0.62 respectively, $p < 0.001$. While there is an apparent relationship between performance on some measures, the Language VAT evaluates language functioning from a conceptually different framework (please refer to Chapter 2).

Table 5.4: Pearson's Correlations between the Language Video Assessment Tool Tests and Other Language Measures (n=30)

	LC	RI	FD	MSS
CASL - Meaning from Context	.69			
CASL - Sentence Expression		.62		
CASL – Grammatical Morphemes				.86
TILLS - Story Retelling		.88		
TILLS – Language Comprehension	.86			
TILLS – Following Directions			.84	

Test-Retest Reliability

This is a factor determined by the variation between scores or different evaluative measurements of the same subject taking the same test during a given period of time. If the test is a strong instrument, this variation would be expected to be low (Osborne, 2008). The Language VAT was administered to 30 randomly selected examinees, ages 5-0 through 12-11 over two testing periods. The interval between the two periods ranged from 16 to 20 days. To reduce recall bias, the examiners did not inform the examinees at the time of the first administration that they would be tested again. All retesting was done by the very same examiners who administrated the test the first time. The results are listed in Table 5.5. The test-retest coefficients for the subtests were all greater than .80 and those for the composite exceeded .90. The size of these coefficients support test-retest reliability of the Language VAT.

Table 5.5

Test - Retest Reliability

		1st Test		2nd Test		Correlation Coefficient
Age Groups	Number	Mean	SD	Mean	SD	
1 , 2, & 3	30					
LC		83	2	82	1	0.91
RI		45	1	46	1	0.88
FD		48	2	48	2	0.84
MSS		56	1	57	1	0.89
4, 5, & 6	30					
LC		94	2	11	1	0.90
RI		52	1	53	1	0.86
FD		53	3	53	2	0.89
MSS		60	2	61	1	0.92
7 & 8	30					
LC		98	2	10	1	0.91
RI		56	2	57	2	0.87
FD		57	1	57	1	0.80
MSS		64	2	65	1	0.93

Abbreviations: LC, Language Comprehension; RI, Restating Information; FD, Following Directions; MSS, Morphology and Sentence Structure

Inter-rater Reliability

Interrater reliability measures the extent to which consistency is demonstrated between different raters with regard to their scoring of examinees on the same instrument (Osborne, 2008). For the Language VAT, inter- rater reliability was evaluated by examining the consistency with which the examiners are able to follow the test scoring procedures. Data was examined by five California licensed speech-language pathologists who independently evaluated 24 test administrations that were selected in a random manner from the normative sample. The raters had one training session during which the item-by-item scoring rules, and the procedures of the study were presented before being asked to score the same verbatim

responses of the 24 randomly selected examinees. The results of the scorings were correlated. The coefficients were averaged using the z-transformation method.

Table 5.6

Standard Errors of Measurement by Age Groups

	Subtests			
	1	2	3	4
Age Group	LC	RI	FD	MSS
7-0 to 7-11	1	2	1	2
8-0 to 8-11	1	2	2	1
9-0 to 9-11	2	2	1	2
10-0 to 10-11	1	2	1	1
11-0 to 11-11	1	2	1	1
12-0 to 13-11	1	2	1	1
14-0 to 15-11	1	2	1	1
16-0 to 18-11	1	1	1	1
Average	1	1	1	1

Abbreviations: LC, Language Comprehension; RI, Restating Information; FD, Following Directions; MSS, Morphology and Sentence Structure

Table 5.7 Inter-rater Reliability Coefficients, Language Video Assessment Tool

Test	*Reliability*
Restating Information (N=29)	*.81*
Following Directions (N=29)	*.79*
Morphology & Sentence Structure (N=29)	*.84*
Language Comprehension (N=29)	*.92*

Effectiveness of Remote Virtual Assessment

Over the past few years, the need for valid and reliable remote assessments has become more evident. In March 2020, we saw many schools and clinics around the world close their doors and turn to virtual speech and language services due to the COVID-19 pandemic. Now, as we are moving our way out of the pandemic, we are continuing to see virtual speech and language services. The reason, possibly, is because virtual speech and language services work (Gabel, Grogan-Johnson, Alvares, Bechstein, & Taylor, 2013) and can be more convenient for some families and individuals.

When we consider the individuals who are receiving speech and language services, the majority are in a critical period of speech and language development (Nicholas & Geers, 2006), and thus, it is crucial that services continue on in order to avoid negative effects on academic performance, peer relationships, and

overall quality of life (Wales, Skinner, & Hayman, 2017; Taylor, Armfield, Dodrill, & Smith, 2014; Kaiser & Roberts, 2011). Previous research has suggested that tele-practice can be an effective model for assessment and treatment (Wales, Skinner, & Hayman, 2017; Keck & Doarn, 2014; Theodoros, 2008; Gabel, Grogan-Johnson, Alvares, Bechstein, & Taylor, 2013). Additionally, the American Speech-Language-Hearing Association (2020) has approved tele-practice as an appropriate method for the assessment and treatment of speech and language disorders. In order to feel confident in the accuracy, reliability, and validity of remote assessments, clinicians can evaluate how scores obtained during remote assessment compare to those scores obtained from in-person administration.

The present study compares speech sound performance results of in-person versus remote administrations of the *Language Video Assessment Tool* (VAT). In order to examine the equivalency between in-person and remote assessments, a test-retest design was used for this study. Each individual who participated in this study was tested twice with the *Language Video Assessment Tool* (VAT), once in-person and once remotely. The same clinician administered both the in-person and remote assessment for each participant. Additionally, the order of which assessment format (in person vs. remote) occurred was counterbalanced. The purpose of the present study is to determine if there are any significant differences in language performance results when testing in-person compared to testing remotely. The present study will also evaluate rater-reliability by evaluating if there are any differences in the clinician's ratings of performance when testing occurs in-person vs. remotely.

The *Lavi Institute* provides a technical manual for the administration and scoring of the *Language VAT.* It is a requirement that the clinician administering the test read and become familiar with the administration, recording, and scoring procedures before using this, or any, assessment tool.

METHOD

Participants

One hundred and six children, aged 5 years, 0 months, to 14 years, 0 months participated in this study. The sample consisted of forty-nine who were considered typically developing and fifty-seven with a previously diagnosed developmental language delay. Demographic characteristics are reviewed in Table 6. The study's sample was balanced for age, gender, and race or ethnic group.

Four examiners participated and administered the assessment used in this study. All examiners were state licensed, ASHA-certified speech-language pathologists (SLPs). The SLPs collected data from September 2020 to December 2022. The SLPs were recruited through *The Lavi Institute*, a research and professional development company. All examiners received compensation for their participation in the study. The one hundred and six participants were also recruited through *the Lavi Institute* and received compensation (e.g., gift card) for their participation.

Materials and Procedures

Prior to all in-person and remote assessments, parent consent was provided to assess each child. Parents also provided consent to have their child's data included for the purpose of this study. Examiners confirmed with parents the day before the remote assessment took place that each child had access to an electronic device, such as a laptop or tablet, with headphones and a built-in microphone. Remote

administration was completed securely over the online Zoom platform. Individual meeting links with passwords were provided for each participant and additional licensing was provided for the examiner to secure HIPAA compliance.

The *Language VAT* is composed of short pre-recorded video segments. Therefore, clinicians used an electronic device during both in-person and remote administrations to access the video-based *Language Video Assessment Tool.*

During remote assessment, the examiner used the screen-sharing feature on Zoom to present and administer the *Language VAT*. After displaying a test item to the student, the examiner paused the test, stopped screen-share, and asked the test item questions per test instructions. The clinician would then listen carefully to the answers. Then, the examiner would start screen-share again and move on to the next item and continue the process until all of the *Language VAT* items were administered.

During each participant's first assessment, he/she was fully assessed using the *Language VAT.* Each participant was then scheduled for his/her follow-up assessment at least three weeks later. A student's speech sound production skills are not expected to change significantly during this time period. Thus, the test-retest method is beneficial in comparing the results of a student's in-person versus remote performance. Additionally, due to this research design, the present study counterbalanced the order of the test format. For example, half of the participants in the typically developing group and half of the participants in the clinical group received an in-person assessment the first time they were assessed and then received remote assessment the second time. The remaining participants received the remote administration the first time they were assessed and an in-person assessment on the second test date.

During both in-person and remote assessments, examiners recorded each participant's responses on the online digital protocol. The results of each assessment were then calculated on the test's website page. The *Language VAT* yields a raw score, standard score, and percentile rank. Participants' standard scores from both testing formats were compared to obtain test-retest reliability. Raw scores from both testing conditions were used to obtain rater-reliability.

RESULTS

Test-retest reliability is the ability for a test to reveal the same score and/or diagnosis when given more than once over a short interval of time. This method was used to determine if the remote administration of the *Language VAT* would reveal the same score and/or diagnosis as the in-person administration. The *Language VAT* was administered twice to one hundred and six participants, aged 5 years, 0 months, to 14 years, 0 months, once in-person and once remotely. The interval between the two testing dates ranged from 20 to 25 days. Participants had the same examiner during the first and second administration. The results are displayed below in Table 1. All participants were grouped initially for primary analysis. The test-retest coefficients for the in-person and remote formats were greater than .80 indicating strong test-retest reliability.

Mean raw scores and standard deviations for in-person and remote standard scores of the *Language VAT* are provided in Table 7. The variance in means across groups is composed of the expected range of performance for typically developing participants (ranging from 5 years, 0 months, to 14 years, 0 months) with the expected range of performance for those with a developmental language delay (ranging from 5 years, 0 months, to 14 years, 0 months). To calculate the effect size, the difference between the mean standard scores of the two testing instances was divided by the pooled standard deviation. An effect size range from 0.02 to 0.16 was realized for the entire sample. An effect size of 0.2

is considered small, 0.5 is considered medium, and 0.8 is considered large (Cohen, 1992). As such, the observed effect sizes were considered small meaning there is insignificant change between the two test conditions (i.e., in-person and remote). Additionally, there were no statistically significant differences found between in-person and remote administrations for the Language VAT.

In order to investigate the reliability of the examiner's ratings, raw scores from in-person and remote testing were compared for each participant. To calculate rater reliability, the intraclass correlation coefficient was used, following the method outlined by Shrout and Fleiss (1979). The intraclass correlation coefficients were .97 for the Language VAT indicating a very high level of agreement across the test administration conditions (i.e., in-person and remote) for the same participant.

DISCUSSION

The purpose of this study was to determine if administering the *Language VAT* remotely would result in the same findings as if it was administered in-person. One hundred and six children students participated in this study and each participant was assessed with the *Language VAT* remotely and in-person. There was an average three-week gap between each test session. Additionally, test order was counterbalanced so that some students received the remote administration first and some received the in-person administration first. Each student's remote and in-person assessment results were compared, and there were no significant differences found between the two formats of assessment. Additionally, remote and in-person assessment resulted in strong reliability of raw and standard scores.

The results of this study demonstrate that in addition to successful in-person administration, the *Language VAT* can also be successfully administered remotely via a secure online platform such as Zoom. Remote assessment does not appear to impact an individual's language comprehension and spoken language performance or the examiner's ability to adequately rate an individual's language comprehension and spoken language production. Additionally, the results of the present study provide evidence that assessment tools can be successfully adapted for remote use and continue to yield valid and reliable results.

In the future, studies can continue to investigate the use of in-person assessment tools adapted for remote administration. Additionally, larger sample sizes with more diverse clinical populations should be used to determine the equivalency of normative assessments via remote administration. In doing so, the findings of future studies can establish whether remote administration of assessments is appropriate. Future studies should also investigate the use of other virtual online platforms and examine if there are any extraneous factors that may impact remote vs. in-person assessment administration. By continuing to investigate the use of remote assessments, clinicians can feel more confident using remote assessments and also guide researchers and test developers in the future.

Table 6

Demographics of the Equivalency Sample

Sample Size = 106

Demographic	*N* Normative Sample	% Normative Sample
Gender		
Male	60	57%
Female	46	43%
Total	106	100%
Race		
White	41	39%
Black	19	18%
Asian	9	8%
Hispanic	32	30%
Other	5	5%
Total	106	100%
Clinical Groups		
	57	54%

Table 7

In-Person vs. Remote Administration Equivalency of Standard Scores, Correlations and Effect Sizes

		In-Person		Remote			
	N	Mean	SD	Mean	SD	*r*	Effect Size
Typically Developing							
Following Directions	49	56	0.8	55	0.5	.98	0.02
Restating Information		55	3.4	53	3.1	.91	0.04
Language Comprehension		95	3.1	96	3.6	.96	0.10
Morphology and Sentence Structure		64	1.2	64	0.9	.93	0.07
Language Impairment							
Following Directions	57	39	2.1	38	3.0	.98	0.03
Restating Information		29	3.7	27	2.5	.93	0.07
Language Comprehension		54	3.5	53	3.1	.97	0.09
Morphology and Sentence Structure		39	2.8	37	2.4	.93	0.16

Chapter

6

Highlights of the Language Video Assessment Tool

The results of the *Language Video Assessment Tool* provide information on the spoken language comprehension and expressive language skills that children and adolescents require to succeed in school and social situations. This assessment is particularly valuable to individuals who have delays in spoken language comprehension, expressive language, language integration, literacy, and social interactions. Data obtained from the *Language VAT* is useful in determining eligibility criteria for a student with a developmental language disorder.

Strong Psychometric Properties

The *Language Video Assessment Tool* was normed on a nationwide standardization sample of 1012 examinees. The sample was stratified to match the most recent U.S. Census data on gender, race/ethnicity, and region. Please refer to Chapter 4 for more information of the standardization process.

The *Language VAT* areas have strong sensitivity and specificity (above 80%), high internal consistency, and test-retest reliabilities. Criterion-related validity studies were conducted during standardization, with over 1012 participants. Please refer to Chapter 5 for more information on the summary results of the reliability and validity studies.

The contextual background and theoretical background sections described in Chapters 1 and 2 provide construct validity of the *Language VAT*. Additionally, please refer to chapter 1 for descriptions of each language skill observed and literature reviews to support this type of measurement included in the *Language VAT*.

Ease and Efficiency of Administration and Scoring

The *Language Video Assessment Tool* consists of four tests: language comprehension, story retell, following directions, and morphology and syntax. The *Language VAT* score converting software is available on the *Video Assessment Tools* website. Please review Chapter 3 for more information on the easy and effective administration process.

The *Language VAT* utilizes visually appealing videos and pictures to keep students engaged and motivated during assessment. The protocols for the *Language VAT* can be scored easily online, and our test converting software works fast to provide you with standard scores and percentiles. Additionally, our report generator will save you time when it comes to completing the write-up portion of your evaluation.

References:

Allen, R. J., & Waterman, A. H. (2015). How does enactment affect the ability to follow instructions in working memory? *Memory and Cognition, 43(3),* 555-561.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A.-M. (2004). A structural analysis of working memory and related cognitive skills in early childhood. *Journal of Experimental Child Psychology, 87,* 85–106.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

American Speech-Language-Hearing Association. (2016). *Definitions of communication disorders and variations*. Retrieved October 26, 2016 from http://www.asha.org/policy/RP1993-00208/.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). Standards for educational and psychological testing. Washington, DC: AERA.

Archibald, L. (2013). The language, working memory, and other cognitive demands of verbal tasks. *Topics in Language Disorders, 33,* 190–207.

Barako Arndt, K., & Schuele, C. M. (2012). Production of infinitival complements by children with specific language impairment. *Clinical Linguistics & Phonetics, 26,* 1–17

Barako Arndt, K., & 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.

Beitchman, J. H, Wilson, B., Johnson, C. J, et al. (2001). Fourteen-year follow-up of speech/language impaired and control children: Psychiatric outcome. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 75–82.

Bergman-Nutley, S., & Klingberg, T. (2014). Effect of working memory training on working memory, arithmetic and following instructions. *Psychological Research, 78(6),* 869-877.

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, 39*(2), 215-237.

Bishop, D. V. M. & McDonald, D. (2009). Identifying language impairment in children: combining language test scores with parental report. *International Journal of Language and Communication Disorders, 44,* 600–615.

Bishop, D. V. M., & Donlan, C. (2005). The role of syntax in encoding and recall of pictorial narratives: Evidence from specific language impairment. *British Journal of Developmental Psychology, 23,* 25–46.

Bloom, L., Tackeff, J., & Lahey, M. (1984). Learning *to* in complement constructions. *Journal of Child Language, 11,* 391–406.

Boudreau, D., & Hedberg, N. (1999). A comparison of early literacy skills in children with specific language impairment and their typically developing peers. *American Journal of Speech-Language Pathology, 8,* 249–260.

Brinton, B., Fujiki, M., & McKee, L. (1998) Negotiation skills of children with specific language impairment. *Journal of Speech, Language and Hearing Research 41,* 927–40.

Burdelski, M., and Evaldsson, A. C. (2019). Young children's multimodal and collaborative tellings in family and preschool interaction. *Research on Children and Social Interaction, 3,* 1–5.

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

Cain, K. & Towse, A. S. (2008). To get hold of the wrong end of the stick: Reasons for poor idiom understanding in children with reading comprehension difficulties. *Journal of Speech, Language, and Hearing Research, 51,* 1538–1549.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (2001). Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation. *Language, Speech, and Hearing Services in Schools, 32*, 38–50.

Catts, H. W., Fey, M., Tomblin, B., Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45,* 1142–1157.

Catts, H.W., Adlof, S. M., & Ellis Weismer, S. (2006). Language deficits of poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*, 278–293.

Clegg, J., Hollis, C., Mahwood, L., & Rutter, M. (2005). Developmental language disorders—a follow-up in later life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry, 46,* 128–149.

Cohen, N. J., Barwick, M., Horodezky, N., Vallance, D. D., & Im, N. (1998). Language, achievement, and cognitive processing in psychiatrically disturbed children with previously identified and unsuspected language impairments. *Journal of Child Psychology and Psychiatry, 36*, 865–78.

Conti-Ramsden, G. & Botting, N. (2004). Social difficulties and victimisation in children with SLI at 11 years of age. *Journal of Speech, Language and Hearing Research, 47*, 145–61.

Conti-Ramsden, G., Durkin, K., Simkin, Z., & Knox, E. (2009). Specific language impairment and school outcomes. I: identifying and explaining variability at the end of compulsory education. International *Journal of Language & Communication Disorders, 44,* 15–35.

Conti-Ramsden, G., Mok, P. L. H., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): strengths and difficulties in social, emotional and behavioral functioning. *Research in Developmental Disabilities*, *34*, 4161–4169.

Craig, H. K. (1993). Social skills of children with specific language impairment: peer relationships. *Language, Speech and Hearing Services in Schools 24*, 206–15.

Craig, H. K & Washington, J. A. (1993). Access behaviors of children with specific language impairment. *Journal of* Speech and Hearing Research 36, 322–37

Crede, M. (2010). Random responding as a threat to the validity of effect size estimates in correlational research. *Educational and Psychological Measurement, 70*, 596–612.

Danner, D. & Rammstedt, B. (2016). Facets of acquiescence: Agreeing with negations is not the same as accepting inconsistency. *Journal of Research in Personality, 65*, 120–129. https://doi.org10.1016/j.jrp.2016.10.010

Diehm E. A., Wood C., Puhlman J., Callendar M. (2020). Young children's narrative retell in response to static and animated stories. *International Journal of Language & Communication Disorders, 55*, 359–372.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge, UK: Cambridge University Press.

Dollaghan, C. A. (2007). The Handbook for Evidence-Based Practice in Communication Disorders. Baltimore, MD: MD Brookes.

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, 56*(8), 848–56.

Durkin, K., Conti-Ramsden, G., & Simkin, Z. (2012). Functional outcomes of adolescents with a history of specific language impairment (SLI) with and without autistic symptomatology. *Journal of Autism and Developmental Disorders*, *42*, 123–138.

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

Eickhoff, J., Betz, S. K., & Ristow, J. (2010). Clinical procedures used by speech-language pathologists to diagnose SLI. *In Symposium on Research in Child Language Disorders,* Madison, WI.

Fazio, B. B., Naremore, R. C., & Connell, P. J. (1996). Tracking children from poverty at risk for specific language impairment. *Journal of Speech, Language, and Hearing Research*, 39, 611–624.

Ferguson, M., Hall, R., Riley, A., Moore, D. (2011). Communication, listening, cognitive, and speech perception skills in children with Auditory Processing Disorder (APD) or Specific Language Impairment (SLI). *Journal of Speech, Language, and Hearing Research, 54,*

211–227.

Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, J. B., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research, 47*(6), 1301–1318.

Florit, E., Roch, M., & Levorato, M. C. (2013). The relation between listening comprehension of text and sentences in preschoolers: Specific or mediated by lower and higher level components? *Applied Psycholinguistics, 34,* 395-415.

Foley, P. B. (2016). Getting Lucky: How Guessing Threatens the Validity of Performance Classifications. *Practical Assessment, Research & Evaluation, 21*(3), 1–23.

Foorman, B., Koon, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). Examining general and specific factors in the dimensionality of oral language and reading in 4th-10th grades. *Journal of Educational Psychology, 107*(3), 884-899.

Fujiki, M., Brinton, B., & Todd, C. M. (1996). Social skills of children with specific language impairment. *Language, Speech and Hearing Services in Schools 27,* 195–202.

Fujiki, M., Brinton, B., & Clarke, D. (2002). Emotion regulation in children with specific language impairment. *Language, Speech, and Hearing Services in Schools 33,* 102–11.

Fujiki, M., Spackman, M. P., Brinton, B., & Illig, T. (2008). Ability of children with language impairment to understand emotion conveyed by prosody in a narrative passage. *International Journal of Language & Communication Disorders, 43*(3), 330–345.

Furr, R. M. & Bacharach, V. R. (2008). *Psychometrics: An introduction*. Thousand Oaks, CA: Sage. Individuals with Disabilities Education Act. (2004). Section 300.8 child with a disability.

Gazella, J., & Stockman, I. J. (2003). Children's story retelling under different modality and task conditions: Implications for standardizing language sample procedures. *American Journal of Speech-Language Pathology, 12,* 61–72.

Gill, C., Moorer-Cook, L., Armstrong, E., & Gill, K. (2012). The ability to follow verbal directions: identifying skill levels and measuring progress. *Canadian Journal of Speech-Language Pathology and Audiology, 36,* 234–247.

Gillam, R. B. & Johnston, J. R. (1992). Spoken and written language relationships in language learning impaired and normally achieving school age children. *Journal of Speech and Hearing Research, 35*(6), 1303-1315.

Gillam, R., McFadden, T. U., & van Kleeck, A. (1995). Improving narrative abilities: Whole language and language skills approaches. In M. E. Fey, J. Windsor, & S. F. Warren (Eds.), *Language intervention: Preschool through the elementary years* (pp. 145–182). Baltimore: Brookes.

Gillam, S. L., Fargo, J. D., & Robertson, K. S. C. (2009). Comprehension of expository text: Insights gained from think-aloud data. *American Journal of Speech-Language Pathology, 18*(1), 82–94.

Goencue, A. & Klein, E. L. (2001). *Children in play, story, and school.* New York, NY: Guilford.

Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, 24, 123–147.

Grove, J., Conti-Ramsden, G., & Donlan, C. (1993). Conversational interaction and decision-making in children with specific language impairment. *European Journal of Disorders of Communication 28,* 141–52.

Guralnick, M. J., Connor, R. T., Hammond, M. A., Gottman, J. M. & Kinnish, K. (1996). The peer relations of preschool children with communication disorders. *Child Development, 67*, 471–89.

Hadley, P. A. & Rice, M. L. (1991) Conversational responsiveness of speech- and language-impaired preschoolers. *Journal of Speech and Hearing Research, 34*, 1308–17.

Hadley, P. (1998). Language sampling protocols for eliciting text-level discourse. *Language, Speech, and Hearing Services in Schools*, *29,* 132–147.

Heilmann, J., Miller, J., Nockerts, A., & Dunaway, C., (2010). Properties of the narrative scoring scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology, 19,* 154-166.

Huang, R., Hopkins, J., & Nippold, M. A. (1997). Satisfaction with standardized language testing. *Language, Speech, and Hearing Services in Schools, 28*(1), 12–29.

Hughes, D., McGillivray, L., & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessment*. Eau Claire, WI: Thinking Publications.

Hulme, C. & Snowling, M. J. (2013). *Developmental disorders of language learning and cognition.* Chichester, UK: Wiley-Blackwell.

Jackson S., Pretti-Frontczak K., Harjusola-Webb S., Grisham-Brown, J., & Romani, J. (2009). Response to intervention: implications for early childhood professions. *Language, Speech and Hearing Services in Schools 40*, 424–434.

Joffe, V. & Black, E. (2012). Social, emotional, and behavioral functioning of secondary school students with low academic and language performance: Perspectives from students, teachers, and parents. *Language, Speech, and Hearing Services in Schools, 43,* 461–473.

Johnson, C. J., Beitchman, J. H., & Brownlie, E. B. (2010). Twenty-year follow-up of children with and without speech– language impairments: family, educational, occupational, and quality of life outcomes. *American Journal of Speech–Language Pathology*, *19*, 51–65.

Juel, C. Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78,* 243-255.

Justice, L. M., Bowles, R. P., Pence Turbull, K. L., & Skibbe, L. E. (2009). School readiness among children with varying histories of language difficulties. *Developmental Psychology*, *45,* 460–476.

Kim, Y.-S. G. (2015). Developmental, component-based model of reading fluency: An investigation of word reading fluency, text reading fluency, and reading comprehension. *Reading Research Quarterly, 50,* 459–481.

Kim, Y.-S., Al Otaiba, S., Wanzek, J., & Gatlin, B. (2015). Towards an understanding of dimension, predictors, and gender gaps in written composition. *Journal of Educational Psychology, 107,* 79-95.

Kim, Y.-S. & Phillips, B. (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly, 49,* 269-281.

Kim, Y.-S., Wagner, R., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: a longitudinal study from grade 1 to grade 2. *Journal of Experimental Child Psychology, 113,* 93-111.

Kim, Y.-S. G., & Wagner, R. K. (2015). Text (Oral) reading fluency as a construct in reading development: An investigation of its mediating role for children from Grades 1 to 4. *Scientific Studies of Reading, 19,* 224-242.

Laing, S. & Kamhi, A. (2002). The use of think-aloud protocols to compare inferencing abilities of average and below-average readers. *Journal of Learning Disabilities, 35,* 437–448.

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics, 32,* 69–92.

Lepola, J., Lynch, J., Laakkonen, E., Silvén, M., & Niemi, P. (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4- to 6-year old children. *Reading Research Quarterly, 47,* 259-282.

Liles, B. Z., Duffy, R. J., Merritt, D. D., & Purcell, S. L. (1995). Measurement of narrative discourse ability in children with language disorders. *Journal of Speech and Hearing Research, 38,* 415–425.

Lindsay, G. & Dockrell, J. (2000). The behaviour and self-esteem of children with specific speech and language difficulties. *British Journal of Educational Psychology, 70,* 583–601.

Lindsay, G., Dockrell, J., Strand, S. (2007). Longitudinal patterns of behaviour problems in children with specific speech and language difficulties: child and contextual factors. *British Journal of Educational Psychology, 77,* 811–28.

Lindsay, G. & Dockrell, J. (2012). Longitudinal patterns of behavioral, emotional, and social difficulties and self-concepts in adolescents with a history of specific language impairment. *Language, Speech, and Hearing Services in Schools, 43,* 445–460.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: A comment on Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review, 109,* 35–54.

Mainela-Arnold, E., Evans, J., & Coady, J. (2008). Lexical representations in children with SLI:

evidence from a frequency- manipulated gating task. *Journal of Speech, Language, and Hearing Research, 51,* 381–393.

Mandler, J., & Johnson, N. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9,* 111–151.

Marinellie, S. A. (2004). Complex syntax used by school- age children with specific language impairment (SLI) in child-adult conversation. *Journal of Communication Disorders, 37,* 517–533.

Marton, K., Abramoff, B., & Rosenzweig, S. (2005). Social cognition and language in children with specific language impairment (SLI). *Journal of Communication Disorders, 38,* 143–62.

Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *American Psychologist,* 63*(8)*, 760-769.

McCauley, R. J., & Strand, E. A. (2008). A review of standardized tests of nonverbal oral and speech motor performance in children. *American Journal of Speech-Language Pathology, 17*(1), 81–91.

McFadden, T. U., & Gillam, R. B. (1996). An examination of the quality of narratives produced by children with language disorders. *Language, Speech, and Hearing Services in Schools, 27,* 48–56.

Miller, J. F. (1981). *Assessing language production in children*. Baltimore, MD: University Park Press.

Miller, J., & Klee, T. (1995). Computational approaches to the analysis of language impairment. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 545–572). Oxford, England: Blackwell.

Mineo, B., Peischl, D. & Pennington, C. (2008). Moving targets: The effect of animation on identification of action word representations. *Augmentative and Alternative Communication, 24,* 162–173.

Mullen, R., & Schooling, T. (2010). The National Outcomes Measurement System for pediatric speech-language pathology. *Language, Speech, and Hearing Services in Schools, 41,* 44-60.

Nation, K. (2005). *Children's reading comprehension difficulties.* In M. J. Snowling & C. Hulme (Eds.), The science of reading (pp. 248-265). Oxford: Blackwell Publishing.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language and Hearing Research, 47.*

Nation, K., & Snowling, M. J. (1998). Semantic processing skills and the development of word recognition: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39,* 85-101.

Neumann, L. (2004). *Video modeling: A visual teaching method for children with autism.* Brandon, FL: Willerik Publishing.

Nippold, M. (2007). *Later language development: School-age children, adolescents, and young adults.* (3rd ed.). Austin, TX: Pro-Ed.

Nippold, M. (2010). *Language sampling with adolescents*. San Diego, CA: Plural.

Nippold, M., Hesketh, L., Duthie, J., & Mansfield, T. (2005). Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research, 48,* 1048–1064.

Norbury, C. F. & Bishop, D. V. M. (2003). Narrative skills of children with communication impairments. *International Journal of Language and Communication Disorders, 38,* 287–313.

Oakhill, J. V. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology, 54,* 31-39.

Oakhill, J., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18,* 443–468.

Oakhill, J. V. & Yuill, N. (1996). *Higher order factors in comprehension disability: processes and remediation.* In C. Cornoldi and J.V. Oakhill (Eds.), Reading comprehension difficulties. Mahwah, NJ: Lawrence Erlbaum Associates.

Osborne, J. W. (Ed.). (2008). *Best practices in quantitative methods.* Los Angeles: Sage Publications.

Owen, A., & Leonard, L. (2006). The production of finite and nonfinite complement clauses by children with specific language impairment and their typically developing peers. *Journal of Speech-Language-Hearing Research, 49,* 548–571.

Quirk, R., Greenbaum, S., Leech, G., & Svartik, J. (1985). *A comprehensive grammar of the English language.* London: Longman.

Parsons, S., Schoon, I., Rush, R., & Law, J. (2011). Long-term outcomes for children with early language problems: beating the odds. *Children and Society*, *25*, 202–214.

Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology, 5*(2), 5–21.

Paulhus, D.L. (2002). Socially desirable responding: The evolution of a construct. In H.I. Braun & D.N. Jackson (Eds.), *The role of constructs in psychological and educational measurement* (pp. 37–48). Mahwah, NJ: Erlbaum.

Pearce, W., McCormack, P., & James, D. (2003). Exploring the boundaries of SLI: Findings from morphosyntactic and story grammar analyses. *Clinical Linguistics & Phonetics, 17,* 325–334.

Peña, E. D., Spaulding, T. J., & Plante, E. (2006). The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. *American Journal of Speech-Language Pathology*, *15*(3), 247–254.

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

Poll, G. H. & Miller, C. A. (2013). Late talking, typical talking, and weak language skills at middle childhood. *Learning and Individual Differences, 26,* 177–84.

Psyridou, M., Eklund, K., Poikkeus, A. M, & Torppa, M. (2018). Reading outcomes of children with delayed early vocabulary: A follow-up from age 2–16. *Research in Developmental Disabilities, 78,* 114–24.

Puranik, C. S., Lombardino, L. J., & Altmann, L. J. (2007). Writing through retellings: An exploratory study of language-impaired and dyslexic populations. *Reading and Writing, 20*(3), 251-272.

Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). Frog, Where are you? Narratives in children with specific language impairment, early focal brain injury and Williams syndrome. *Brain & Language, 88,* 229–247.

Rescorla, L. (2009). Age 17 language and reading outcomes in late-talking toddlers: support for a dimensional perspective on language delay. *Journal of Speech Language and Hearing Research, 52*(1), 16–30.

Robertson, E. K., & Joanisse, M. F. (2010). Spoken sentence comprehension in children with dyslexia and language impairment: The roles of syntax and working memory. *Applied Psycholinguistics, 31,* 141-165.

Rumelhart, D. (1975). Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211–236). New York, NY: Academic Press.

Schlosser, R. W., Koul, R., Shane, H., Sorce, J., Brock, K., Harmon, A., Moerlein, D., & Hearn, E. (2014). Effects of animation on naming and identification across two graphic symbol sets representing verbs and prepositions. *Journal of Speech Language Hearing Research, 57,* 1–13.

Schuele, C., & Tolbert, L. (2001). Omissions of obligatory relative markers in children with specific language impairment. *Clinical Linguistics & Phonetics, 15,* 257–274.

Schuele, M., & Dykes, J. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment. *Clinical Linguistics and Phonetics, 19,* 295 – 318.

Scott, C. M., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research*, 43, 324–339.

Scott, C., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research, 43,* 324–339.

Scott, C. M. & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech Language and Hearing Research, 43*(2), 324-339.

Seiger-Gardner, L., Schwartz, R. (2008). Lexical access during word production in children with and without SLI- A cross-model picture-word interference study. *International Journal of Language and Communication Disorders, 43,* 528–551.

Sherer M., Pierce, K. L., Paredes, S., Kisacky, K. L., Ingersoll, B., & Schreibman, L. (2001). Enhancing conversation skills in children with autism via video technology. *Behavior Modification, 25,* 140-158.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Snowling, M. J. & Hulme, C. (2012). Annual Research Review: The nature and classification of reading disorders—a commentary on proposals for DSM-5. *Journal of Child Psychology and Psychiatry, 53,* 593–607.

Spaulding, T.J., Plante, E., & Farinella, K.A. (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 behavioural, emotional and social difficulties in individuals with a history of specific language impairment (SLI). *Journal of Communication Disorders, 44,* 186–199.

Stein, N., & Glenn, C. (1979). An analysis for story comprehension in elementary school. In R. Freedle (Ed.), *New directions in discourse processing* (Vol. 2, pp. 53–120). Norwood, NJ: Ablex.

Sweeting, H. & West, P. (2001). Being different: Correlates of the experience of teasing and bullying at age 11. *Research Papers in Education, 16*(3), 225–246.

Takagi, T. (2019). Referring to past actions in caregiver–child interaction in Japanese. *Research on Children and Social Interaction, 3,* 92–118.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*(6), 1245–1260.

Vallance, D. D, Im, N., & Cohen, N. J. (1999). Discourse deficits associated with psychiatric disorders and with language impairments in children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40,* 693–705.

van den Broek, P. (1997). Discovering the cement of the universe: The development of event comprehension from childhood to adulthood. In P. van den Broek, P. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension: Bridging fictional and actual*

events (pp. 321–342). Mahwah, NJ: Erlbaum.

Verhallen, M. J., Bus, A. G., & De Jong, M. T. (2006). The promise of multimedia stories for kindergarten children at risk. *Journal of Educational Psychology, 98*, 410–419.

Westby, C. E., Stevens-Dominguez, M., & Oetter, P. (1996). A performance/competence model of observational assessment. *Language, Speech, and Hearing in the Schools, 27*, 144–156.

Westby, C. (2005). Assessing and facilitating text comprehension problems. In H. Catts & A. Kamhi (Eds.), *Language and reading disabilities* (pp. 157–232). Boston, MA: Allyn & Bacon.

Wetzel, E., Lüdtke, O., Zettler, I., & Bohnke, J. R. (2016). The stability of extreme response style and acquiescence over 8 years. *Assessment, 23*, 279–291.

Windsor, J., Scott, C. M., & Street, C. K. (2000). Verb and noun morphology the spoken and written language of children with language learning disabilities. *Journal of Speech, Language and Hearing Research, 43,* 1322-1336.

Yawkey, T. D., Aronin, E. L., Streett, M. A., & Hinojosa, O. M. (1974). Teaching oral language to young Mexican-Americans. *Elementary English, 51*(2), 198–202.

Zielinski, B. W., Bench, R. J., & Madsen, M. F. (1997). A follow-up study of the later reading comprehension ability of language-deficient preschoolers who recovered before starting school. *Asia Pacific Journal of Speech, Language and Hearing, 2*(2), 111–23.

Zumbo, B. D. & Chan, E. K. (2014). *Validity and Validation in Social, Behavioral, and Health Sciences*. Springer.