Reading Comprehension Listening Comprehension in Individuals with Down Syndrome and Word Reading-Matched Typically Developing Children

Key words: Down syndrome, listening comprehension, reading comprehension, literacy

Alison Prahl, PhD[a]; alison_hessling@baylor.edu

Baylor University

[a]One Bear Place #97332, Waco, TX 76798, USA

C. Melanie Schuele, PhD[b]; melanie.schuele@vumc.org

Vanderbilt University Medical Center

[b]1215 21st Ave S, 8310 MCE, South Tower, Nashville, TN 37232, USA

Author Note

Alison Prahl, PhD, Assistant Professor, Communication Sciences and Disorders Department, Baylor University, Waco, TX; C. Melanie Schuele, PhD, Associate Professor, Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN.

This research, completed as part of Alison Hessling Prahl's dissertation at Vanderbilt University, was supported by the US Department of Education Preparation of Leadership Personnel grant (H325D140087) and in part by Vanderbilt CTSA award No. UL1TR000445 from the National Center for Advancing Translational Sciences and a Russell G. Hamilton Graduate Leadership Institute Dissertation Enhancement Grant, Vanderbilt University Graduate School. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US Department of Education, the National Center for Advancing

Translational Sciences, or the Russell G. Hamilton Graduate Leadership Institute. Both authors read and approved the final manuscript. The authors have no conflict of interest to disclose.
Correspondence concerning this article should be addressed to Alison Prahl, alison_prahl@baylor.edu, 254-710-2572, or One Bear Place #97332, Waco, TX 76798.

Abstract

Purpose: The purpose of this study was to explore the reading comprehension and listening comprehension performance of English-speaking children with Down syndrome (DS) compared with word reading-matched typically-developing (TD) children.

Method: Participants included 19 individuals with DS (mean age = 17;2, Range = 11;1-22;9) and 19 word reading-matched TD children (mean age = 7;2, Range = 6;6-8;1). Participants completed three norm-referenced measures of reading comprehension and three norm-referenced measures of listening comprehension. Dependent variables were raw scores on each measure, with the exception of scaled scores on one reading comprehension measure.

Results: Independent-samples t-tests with Bonferroni adjusted alpha levels of .008 revealed a significant between-group difference for two of three reading comprehension measures. The mean raw scores were lower for the DS group than the TD group, with large effect sizes. Independent-samples t-tests with Bonferroni adjusted alpha levels of .008 revealed a significant between-group difference for three of three listening comprehension measures. The mean raw scores on the three measures were lower for the DS group than the TD group, with large effect sizes.

Conclusion: The DS group, despite being matched on word reading to the TD group, demonstrated reduced reading comprehension skills as compared with the TD group. Thus, as individuals with DS acquire word reading skills, it appears that they are unable to translate word reading success to achieve reading comprehension at the expected level (i.e., as indexed by typical readers). The between-group differences in listening comprehension suggest that deficits in listening comprehension likely are a barrier to reading comprehension proficiency for children with DS. Listening comprehension may be a malleable factor that can be targeted to improve reading comprehension outcomes for individuals with DS.

Reading Comprehension Listening Comprehension in Individuals with Down Syndrome and Word Reading-Matched Typically Developing Children

Historically children with intellectual disabilities (ID), including children with Down syndrome (DS), were not expected to acquire much more than rudimentary reading skill (e.g., limited sight word recognition). Consequently, reading instruction was quite limited. In recent years, researchers have focused on developing and validating approaches to improve reading outcomes for children with ID and results argue for consistent and systematic efforts to teach children with ID to read (e.g., Allor et al., 2010; Allor et al., 2013; Lemons et al., 2017). Importantly, if a child with ID achieved only second grade reading skill, for example, (s)he would be able to read for pleasure and to use print in vocational settings to facilitate work achievement and overall independence. With respect to children with DS, Lemons and colleagues (2015) argued that intervention protocols that consider the phenotypic characteristics of DS are more likely to result in change than those that do not. The characteristics to consider include cognition and short-term memory deficits, speech and language deficits, behavioral challenges related to increased task demands, as well as relative strengths in visual processing and language comprehension, including vocabulary. Because children with DS have receptive and expressive language skills that are discrepant from their broad cognitive profile, it is imperative to target not only word decoding and recognition skills but reading comprehension skills as well. Language deficits provide good reason to expect that a child with DS who can accurately read a passage (i.e., accurate oral reading) will not achieve expected comprehension of the passage. The purpose of this study was to explore the reading comprehension and listening comprehension of children with DS compared with word reading-matched typically-developing (TD) children.

Reading Comprehension Theoretical Framework

Reading comprehension, the desired outcome of a person's efforts to decode printed text, is a multidimensional construct (Kamhi & Catts, 1999; Catts & Kamhi, 2017; Elleman & Compton, 2017; Fuchs et al, 2018). As such, successful reading comprehension requires the coordination of multiple underlying cognitive and linguistic processes (Kintsch, 1998; Snow, 2002; Elleman & Compton, 2017). In the Simple View of Reading, reading comprehension is hypothesized as the product of word decoding and linguistic comprehension (Gough & Tunmer, 1986). This Simple View recognizes that translation of orthographic units (i.e., words) to phonological equivalents is insufficient to yield text comprehension. What is also necessary is the integration of word recognition with processes of language comprehension. In their seminal publication, Gough and Tunmer (1986) defined decoding as context-free word recognition, operationalized as a measure of decoding pseudo words (e.g., *stenk*). They defined linguistic comprehension as the interpretation of lexical information, sentences, and discourse, operationalized as a measure of listening comprehension. Importantly, the Simple View presumes that, once printed text is decoded, the reader applies to the text the same mechanisms or processes which the reader would bring to comprehending its spoken equivalent. Thus, reading comprehension closely parallels linguistic comprehension within the Simple View of Reading. In this study, we operationally defined listening comprehension (as an indicator of linguistic comprehension) as constructing meaning from read-aloud written text (i.e., another person reads the text aloud) and reading comprehension as constructing meaning from one's independent decoding of written text.

Simple View of Reading in Down Syndrome

Only a few studies of individuals with DS have evaluated the components of the Simple View of Reading—word reading, listening comprehension, and reading comprehension (Boudreau, 2002; Laws et al., 2016; Nash & Heath, 2011; Roch & Levorato, 2009). Boudreau (2002) compared the literacy skills of 20 children and adolescents with DS (5- to 17-year-olds) to 20 nonverbal mental age-matched TD peers (3- to 5-year-olds). There was a significant between-group difference on word reading (d = 1.07), with the DS group outperforming the mental age-matched group. However, there was no between-group difference on reading comprehension (d = .39). Thus, despite an advantage in word reading, the children with DS did not demonstrate a comparable advantage in reading comprehension. By applying the Simple View to Boudreau's results, we hypothesize that children with DS will have greater difficulties in listening comprehension as compared with their mental age-matched peers.

A similar pattern of results emerged in Nash and Heath (2011) who compared children with DS (n = 13; 11- to 19-year-olds) to word reading-matched typical readers (n = 13; 8- to 9-year-olds). There was a significant between-group difference on reading comprehension, with the word reading-matched group outperforming the DS group (*d* = 2.34). The DS group did not achieve the expected level of reading comprehension for their word reading achievement level. Nash and Heath also compared the DS group to reading comprehension-matched typical readers (n = 13; 7- to 9-year-olds). There was a significant between-group difference on word reading, with the DS group outperforming the reading comprehension-matched group (*d* = 1.11). The DS group needed greater word reading proficiency than typical readers to achieve comparable reading comprehension proficiency. These findings provide further suggestion of listening comprehension deficits as a source of reduced reading comprehension achievement.

Only Laws et al. (2016) and Roch and Levorato (2009) measured word reading, listening comprehension, and reading comprehension in children with DS, an approach essential to exploring the relation between the three components of the Simple View. Laws et al. (2016) assessed listening and reading comprehension using parallel tasks; the examiner or the child, respectively, read a short sentence and the child selected the picture (from a field of four) that illustrated the sentence. Notably, both tasks only involved sentence-level comprehension, providing a very limited view of listening and reading comprehension. Performance on these tasks was compared in children with DS (n = 14; 6- to 13-year-olds) and word reading-matched typical readers (n = 23; 6- to 8-year-olds). Consistent with Nash and Heath (2011), there was a significant between-group difference on reading comprehension (*d* = 1.69), with the word reading-matched group outperforming the DS group. Despite the simplicity of the task, there was also a between-group difference on listening comprehension (*d* = 2.79), with the word reading-matched group mean exceeding the DS group mean.

Roch and Levorato (2009) matched two groups on reading comprehension, Italian-speaking adolescents with DS (n = 23; 11- to 18-year-olds) and typical readers (n = 23; 6- to 7-year-olds). Consistent with Nash and Heath (2011), there was a significant between-group difference on word reading (*d* = 1.41). They evaluated listening comprehension with a measure parallel to their reading comprehension matching measure; participants answered multiple-choice comprehension questions after listening to a story. The groups differed on listening comprehension (*d* = 1.21) with the between-group difference favoring the typical readers. Thus, when matched on reading comprehension, the DS group had stronger word reading proficiency than the TD group but weaker listening comprehension.

In summary, the extant literature provides a direction for further evaluation of the Simple View in children with DS. The findings of Boudreau (2002) and Nash and Heath (2011) motivate investigation of listening comprehension in children with DS to better understand reading comprehension outcomes. Roch and Levorato's (2009) findings are intriguing but, because Italian has a transparent orthography, the generalization of their findings to English-speaking children with DS is not straightforward. In a language with a transparent orthography, listening comprehension has been shown to be a more powerful predictor of reading comprehension than word decoding (Megherbi et al., 2006). Laws et al. (2016) provided a limited perspective on the relation of listening and reading comprehension because only sentence-level comprehension was tapped. Further, their picture selection task has limited ecological validity as compared with reading comprehension demands in academic or vocational settings.

Additionally, the age range (6- to-13-year-olds) of participants with DS in Laws et al. (2016) may have provided a limited view. As expected, children with DS become readers at older ages than TD children (Laws & Gunn, 2002) and even among older participants, heterogenous reading outcomes are common (Boudreau, 2002). Thus, when the DS participant sample is young, there are potential reading floor effects. Therefore, in our study we evaluated reading comprehension and listening comprehension in individuals with DS from a wider and older age range than Laws et al., with the goal of including more participants who may have developed as readers at a later age. Lastly, previous research has not adequately described whether participants were required to be able to read words at a minimal level (e.g., Boudreau, 2002; Laws et al., 2016). Without exclusionary criteria for word reading, floor effects on reading measures may have been observed due to participants not being able to complete the task rather than reflecting poor comprehension per se (e.g., see Boudreau, 2002 for discussion of skewed distribution and floor effects). Roch and Levorato (2009), the only study that included sufficient reading exclusionary criteria details to allow for replication, implemented a two-phase testing procedure in which participants were excluded if (1) they could not read a short, preschool level story and (2) they scored at or near chance level on a first-grade reading comprehension test. In the current study, we evaluated reading comprehension and listening comprehension with ecologically-valid measures among English-speaking individuals with DS who met a minimum word reading criterion level.

Simple View of Reading: Measurement Method Implications

Given that reading comprehension and listening comprehension are multidimensional constructs, there are multiple and varied ways in which reading comprehension and listening comprehension is measured. How a research team operationally defines and subsequently measures these constructs may influence conclusions drawn. The degree to which measures tap various cognitive and linguistic processes may have important implications for individuals with DS who present with a unique pattern of phenotypic strengths and weaknesses (e.g., Chapman & Hesketh, 2000; Cutting & Scarborough, 2006). Given the myriad measurement methods for reading comprehension and listening comprehension, we evaluated these two constructs across three measurement methods—nonverbal response, cloze procedure, and open-ended questions—selected based on commonly used measures and the DS phenotype (see Table 1). We included individuals with DS ages 11 to 22 years and word reading-matched TD children. We hypothesized that group differences would hold across the various measurement methods, although perhaps the magnitude of the group difference would vary by measurement method. Findings provide unique contributions to the existing, yet limited, literature base that can inform the design of interventions to improve reading outcomes in DS.

The selected methods displayed in Table 1 represent a range of text and response formats that take the DS phenotype into consideration. For instance, because determining the accuracy of verbal responses can be confounded by speech unintelligibility in individuals with DS, the nonverbal response method eliminated the need for a verbal response. Although the remaining methods do not minimize verbal demands, they represent methods of engaging in reading comprehension and listening comprehension that may frequently be encountered in academic and vocational settings. The specific measures were selected because the initial test items at lower levels of difficulty and complexity and the amount of scaffolding provided (i.e., illustrated items on the Woodcock Reading Mastery Test-III Passage Comprehension subtest) were expected to reduce task demands to minimize floor effects. Thus, the selected measures incorporate design elements that align with strengths in the DS phenotype (e.g., visual supports align with strengths in visual processing).

Table 1

Methods of Measuring Listening Comprehension and Reading Comprehension

Method	Text format	Response format	Reading comprehension measure	Listening comprehension measure
Nonverbal response	Phrase and sentence	Nonverbal (pointing, acting out)	KABC Reading/ Understanding subtest	WJ IV Test of Oral Language Understanding Directions subtest
Cloze procedure	Sentence and paragraph	Verbal, one word	WRMT-III Passage Comprehension subtest	WJ IV Test of Oral Language Oral Comprehension subtest
Passage-level with open-ended questions	Paragraph	Verbal	WIAT-III Reading Comprehension subtest	WIAT-III Listening Comprehension subtest

Note. KABC = Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983); WJ IV = Woodcock-Johnson IV (Schrank et al., 2014); WRMT-III = Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); WIAT-III = Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009).

Two research questions were addressed: (1) What is the relation between word reading, listening comprehension, and reading comprehension in individuals with DS? Does this relation differ when compared with TD children? (2) Do individuals with DS differ from children matched for word reading on measures of reading comprehension and listening comprehension?

Method

The study methods were approved by the Vanderbilt University Institutional Review Board.

Participants

Two groups who were matched on word reading participated: (a) individuals with DS and (b) TD children (control group). To form the TD control group, each participant with DS was matched to one TD participant (i.e., a TD participant could only be paired with a single DS participant) based on word-level reading and sex when possible. A TD child was considered an eligible match if his or her raw score on the Word Identification subtest of the Woodcock Reading Mastery Tests-III (WRMT-III; Woodcock, 2011) was within three points of the raw score for a participant with DS. For example, a TD child with a raw score between 7 and 13 could serve as a match for a participant with DS with a Word Identification raw score of 10. The groups were well matched based on word reading, with analyses yielding *p*-values greater than .50 (Mervis & Robinson, 2003; see Table 2).

Table 2

Participant Characteristics Reported as Raw Score Means, Standard Deviations, and Ranges

	DS Group (n = 19)			TD Group (n = 19)			
	Mean	SD	Range	Mean	SD	Range	*p*
Age (months)	206.63	41.68	133–273	86.11	6.54	78–97	.000*
KBIT-2	16.21	5.02	10–28	25.42	5.63	14–34	.000*
ROWPVT-4	77.58	27.85	22–132	101.47	8.71	82–117	.001*
EOWPVT-4	82.95	19.40	50–117	96.79	14.32	68–122	.017*
TACL-4 Grammatical Morphemes	35.53	8.73	19–54	48.16	4.71	41–54	.000*
WRMT-III Word Identification	21.32	6.79	12–37	20.84	6.90	11–34	.832
Arizona-4	88.92	7.27	74–100	97.90	3.34	88–100	.000*

Note. DS = Down syndrome; TD = Typically developing; SD = Standard deviation; KBIT-2 = Kaufman Brief Intelligence Test-Second Edition (Kaufman, 2004); ROWPVT-4 = Receptive One Word Picture Vocabulary Test-Fourth Edition (Martin & Brownell, 2011a); EOWPVT-4 = Expressive One Word Picture Vocabulary Test-Fourth Edition (Martin & Brownell, 2011); TACL-4 = Test of Auditory Comprehension of Language-Fourth Edition (Carrow-Woolfolk & Allen, 2014); WRMT-III = Woodcock Reading Mastery Tests-Third Edition (Woodcock, 2011); Arizona-4 = Arizona Articulation Phonology Scale-Fourth Edition (Fudala & Stegall, 2017).

The eligibility criteria detailed below assisted in identifying participants who were able to participate in study procedures (e.g., adequate hearing and vision, ability to attend to assessments, demonstrated requisite word-level reading skills). The criteria also ensured inclusion of a representative sample of the clinical population of interest. Participation was limited to monolingual English speakers to ensure that performance on language and reading measures was not influenced by exposure to multiple languages. We expect the process of learning to read to differ for children who are learning multiple languages (e.g., Avalos et al., 2007; Fitzgerald, 2003).

DS Group. Nineteen individuals with DS, 11 to 22 years of age (*M* = 17;3, *SD* = 3;6), participated. See Table 3 for participant demographic information. Participants were recruited by distributing study flyers (a) at private schools in Nashville TN and Dallas/Fort Worth TX, (b) with DS community organizations (e.g., Down Syndrome Association of Middle Tennessee, University of Alabama Intellectual Disabilities Participant Registry), (c) on research listservs, and (d) to families whose children had participated in previous research studies in the lab. Individuals were eligible to participate if they (a) had been diagnosed with DS by a physician per parent report, (b) were monolingual English speakers and used spoken language as a primary form of communication, (c) successfully completed the screening battery (i.e., listened to directions, completed assessments), and (d) had normal or corrected-to-normal vision per parent report. Hearing status inclusionary criteria was not used for the DS group to ensure inclusion of a representative sample of participants with DS, who frequently present with mild to moderate hearing loss (Roizen et al., 1993). DS participant hearing screening thresholds are reported (see Table 4). Exclusionary criteria included: (a) correctly reading fewer than 80% of words on the Phonological Awareness Literacy Screening-Kindergarten primer list (PALS-K; Invernizzi et al., 1997) or (b) uncontrolled seizures per parent report. Seven consented individuals with DS were not eligible to participate; one individual did not successfully complete the screening battery, and six individuals did not meet the word reading exclusionary criteria.

Table 3

Participant Demographic Information

	DS Group (n = 19)	TD Group (n = 19)
Sex		
Male	8	6
Female	11	13
Race		
American Indian/Alaska Native	0	0
Asian	0	0
Black/African American	1	1
Native Hawaiian/Other Pacific Islander	0	0
White	17	15
Multiple races	1	1
Not reported	0	2
Ethnicity		
Hispanic or Latino	1	3
Not Hispanic or Latino	17	15
Not reported	1	1
Mother's education level		
Some high school	0	0
High school diploma/GED	1	0
Some college	2	3
Associate's degree	3	0
Bachelor's degree	6	9
Master's degree	5	4
Professional degree	2	3

Note. DS = Down syndrome, TD = typically developing

Table 4

Hearing Screening Thresholds in Decibels (dB) for DS Participants (n = 19)

	Mean	*SD*	Range
Right ear			
500 Hz	36.47	9.48	30–60
1000 Hz	35.59	9.98	30–60
2000 Hz	32.65	6.40	30–55
4000 Hz	39.71	13.17	30–70
Left ear			
500 Hz	35.29	10.07	30–70
1000 Hz	33.24	6.83	30–50
2000 Hz	32.06	6.14	30–55
4000 Hz	32.94	7.92	30–60

Note. dB = Decibels; *SD* = Standard deviation. Two DS participants wore bilateral hearing aids which parents reported to be tested regularly and in working conditioning at the time of the study sessions. We use the term ‘hearing screening thresholds’ because no participant was presented stimuli less than 30dB.

TD Group. Nineteen TD children, 6 to 8 years of age (*M* = 7;2, *SD* = 0;6), participated. See Table 3 for participant demographic information. Participants were recruited from the Nashville TN metropolitan area by distributing study flyers (a) on research listservs (e.g., ResearchMatch, Vanderbilt Research Distribution and Notification List), (b) to families whose children had participated in previous research studies in the lab, (c) to community organizations (e.g., community center, public library), and (d) to families of local elementary school first and second grade students who were reading on grade level. TD children were eligible to participate if they (a) demonstrated oral language skills within normal limits and neurotypical development per parent report, (b) were monolingual English speakers, (c) successfully completed the screening battery (i.e., listened to directions, completed assessments), (d) passed hearing screening in at least one ear, unaided using ASHA standards (ASHA, 2021), and (e) had normal or corrected-to-normal vision per parent report. Exclusionary criteria included: (a) correctly reading fewer than 80% of words on the PALS-K or (b) more than 1.5 standard deviation below the normative mean on the measure of nonverbal cognition. The exclusionary criteria ensured that participants in the control group minimally met the criteria of nonverbal intelligence no more than 1.5 standard deviations below the mean to exclude children with below average intelligence. Five consented TD children were not eligible to participate; four did not meet the word reading exclusionary criteria and two were not monolingual English speakers. Ten consented TD children completed the study procedures but were not included as a participant because they did not match a DS participant on the word reading criteria or did not match on sex.

Procedures

Participants completed two individual sessions (eligibility and assessment) at the university lab, school, community location (e.g., public library), or in their home. See Table 5. All sessions were audio and video recorded. To determine participant eligibility and ascertain whether an individual could comply with assessment procedures, minimally invasive eligibility and descriptive assessments were administered at the initial session (see Tables 2 and 6). Assessments were administered by (a) the first author, a graduate student and certified speech-language pathologist, or (b) a graduate research assistant under the supervision of the first author. Standardized assessments were administered according to the manualized instructions. For each measure, we report the results of a between-group t-test in Table 2. Significant between-group differences were observed on all descriptive measures except word level reading, the matching criteria. Table 6 parallels Table 2 but reports the mean standard scores. Because the age range of the DS group extended beyond the TACL-4 age range (normed for individuals ages 3 to 12), TACL-4 standard scores are not reported for the DS group.

Table 5

Eligibility, Descriptive, and Dependent Variable Measures and Schedule

Construct	Measure(s)	Schedule
Eligibility Assessments		
Hearing screening	Pure tone audiometry	Eligibility session
Word level reading screening	PALS-K primer list	Eligibility session
Nonverbal cognition	KBIT-2 Matrices subtest	Eligibility session
Descriptive Measures		
Receptive vocabulary	ROWPVT-4	Eligibility session
Expressive vocabulary	EOWPVT-4	Eligibility session
Grammar comprehension	TACL-4 Grammatical Morphemes subtest	Eligibility session
Word level reading	WRMT-III Word Identification Subtest	Eligibility session
Speech accuracy	Arizona-4	Eligibility session
Dependent Variable Measures		
Reading comprehension	KABC Reading/ Understanding subtest WRMT-III Passage Comprehension subtest WIAT-III Reading Comprehension subtest	Assessment Session
Listening comprehension	WJ IV Test of Oral Language Understanding Directions subtest WJ IV Test of Oral Language Oral Comprehension subtest WIAT-III Listening Comprehension subtest	Assessment Session

Note. PALS-K = Phonological Awareness Literacy Screening-Kindergarten (Invernizzi et al., 1997); KBIT-2 = Kaufman Brief Intelligence Test-Second Edition (Kaufman, 2004); ROWPVT-4 = Receptive One Word Picture Vocabulary Test-Fourth Edition (Martin & Brownell, 2011a); EOWPVT-4 = Expressive One Word Picture Vocabulary Test-Fourth Edition (Martin & Brownell, 2011); TACL-4 = Test of Auditory Comprehension of Language-Fourth Edition (Carrow-Woolfolk & Allen, 2014); WRMT-III = Woodcock Reading Mastery Tests-Third Edition (Woodcock, 2011); Arizona-4 = Arizona Articulation Phonology Scale-Fourth Edition (Fudala & Stegall, 2017); KABC = Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983); WIAT-III = Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009); WJ IV = Woodcock-Johnson IV (Schrank et al., 2014).

Table 6

Participant Characteristics in Standard Score or Scaled Score Means, Standard Deviations, and Ranges

	DS Group (n = 19)			TD Group (n = 19)		
	Mean	SD	Range	Mean	SD	Range
KBIT-2	52.37	12.25	40–80	109.47	13.26	82–127
ROWPVT-4	59.63	7.87	55–81	112.89	7.80	96–127
EOWPVT-4	62.67	10.81	55–86	111.32	14.57	85–131
TACL-4 Grammatical Morphemes*				11.58	2.22	8–15
TACL-4 GM Age Equivalents	5;8	1;10	3;6–11;9	8;11	2;0	6;3–11;9
WRMT-III Word Identification	61.68	11.07	55–86	110.21	15.91	75–138
WRMT-3 Word ID Grade Equivalents	3.0	2.2	1.3–9.7	3.6	1.8	1.3–7.7
AAPS-4	57.5	15.82	50–96	99.58	1.16	96–100

Note. DS = Down syndrome; TD = Typically developing; SD = Standard deviation; KBIT-2 = Kaufman Brief Intelligence Test-Second Edition (Kaufman, 2004); ROWPVT-4 = Receptive One Word Picture Vocabulary Test-Fourth Edition (Martin & Brownell, 2011a); EOWPVT-4 = Expressive One Word Picture Vocabulary Test-Fourth Edition (Martin & Brownell, 2011); TACL-4 = Test of Auditory Comprehension of Language-Fourth Edition (Carrow-Woolfolk & Allen, 2014); *TACL-4 Scaled scores not reported for DS Group because the age range of the DS group extended beyond the TACL-4 normative age; WRMT-III = Woodcock Reading Mastery Tests-Third Edition (Woodcock, 2011); Arizona-4 = Arizona Articulation Phonology Scale-Fourth Edition (Fudala & Stegall, 2017).

Eligibility Session. Written consent was provided by either a parent or guardian or a person with power of attorney for participants 18 and older. Each participant provided written assent. Each participant's parent/guardian or power of attorney provided demographic background information by completing an intake questionnaire. All eligibility session measures were administered in the same fixed order. The eligibility session lasted 45-60 minutes and thus breaks were not needed.

Hearing screening. Pure tone audiometry with a standard hand-raising response was used to screen hearing acuity in both ears at frequencies of 500, 1000, 2000, and 4000 Hz at 30dB. For the DS group, when a participant failed to respond to a particular frequency at 30 dB, the intensity of the tone was increased until a reliable response was obtained. The highest intensity necessary to elicit a passing response (two out of three presentations) was recorded (see Table 3).

Nonverbal intelligence. The Kaufman Brief Intelligence-Second Edition Matrices subtest (KBIT-2; Kaufman, 2004) was administered as a measure of nonverbal intelligence. Test takers infer a relation or rule in a set of pictures or patterns and point to the picture or pattern that best fits the relation or rule. The KBIT-2 includes simple oral instructions and only requires test takers to answer with a meaningful gesture such as pointing. The K-BIT is normed for individuals ages 4 to 90 and is ideal for those with limited language ability. The mean internal-consistency reliability by age was .88 and the mean test-retest reliability by age was .83, as reported in the K-BIT manual.

Oral language. The Receptive and Expressive One Word Picture Vocabulary Tests-Fourth Editions (ROWPVT-4 and EOWPVT-4; Martin & Brownell, 2011b; Martin & Brownell, 2011a) were administered as measures of receptive and expressive semantic knowledge. For the ROWPVT-4, test takers point to the picture (out of a field of four) that corresponds with the word the examiner says aloud. The ROWPVT-4 manual reported median internal consistency reliability coefficient by age of 0.97 and the test-retest reliability coefficient of .97. For the

EOWPVT-4, test takers name pictures. The EOWPVT-4 manual reported median internal consistency reliability coefficient by age of 0.95 and the test-retest reliability coefficient of .98. These measures are normed for individuals ages 2 to 70 and were found to have strong evidence of content, construct, and criterion-related validity. The Test of Auditory Comprehension of Language-Fourth Edition Grammatical Morphemes subtest (TACL-4; Carrow-Woolfolk & Allen, 2014) was administered as a measure of grammar comprehension. Test takers point to the picture (out of a field of three) that corresponds to stimuli of increasing grammatical complexity presented orally by the examiner. The TACL-4 is normed for individuals ages 3 to 12. Due to limited grammar comprehension characteristic of the DS phenotype, participants with DS did not reach ceiling levels on this measure despite that the DS participant age range extended beyond the normative age range. The TACL-4 manual reported Grammatical Morphemes mean internal consistency reliability of .95 and test-retest reliability of .71. The TACL-4 is a valid measure of oral language based on strong evidence of content-description, criterion-prediction, and construct-identification validity.

Word-level reading. On the PALS-K primer list (eligibility measure), test takers read a list of 20 isolated, real words. Each word read accurately via decoding or automatic recognition is scored as correct; percent correct was calculated. On the WRMT-III Word Identification subtest test takers read isolated, real words. A word is scored correct if read accurately within approximately five seconds, whether it is decoded or automatically recognized. Each DS participant began reading at one of the first three entry points depending on the ease with which they read the PALS-K words and each TD participant began reading at their respective grade level entry point. The manualized instructions were then followed to establish the basal and ceiling. The manual reported mean internal-consistency reliability by school-level socioeconomic status of .93 and the mean test-retest reliability by age of .92. In addition to participant matching, the WRMT-III Word Identification raw scores and standard scores are reported for descriptive purposes. The WRMT-III is normed for individuals ages 4;6 to 79. The manual reported mean split-half reliability coefficient by age of .93 and the test-retest reliability coefficient of .95 for pre-kindergarten through Grade 2, .90 for Grades 3-8, and .88 for Grades 9-12.

Speech. The Arizona Articulation and Phonology Scale-Fourth Revision (Arizona-4; Fudala & Stegall, 2017) was administered as a measure of speech accuracy. Test takers label pictures. If the child does not provide the intended label, the label is modeled by the examiner and repeated by the test taker. The examiner notes speech sound production errors. The Word Articulation Total Score was calculated based on the weighted values (a reflection of how frequently the sound occurs in American speech) of the sounds that were produced accurately. The Arizona-4 is normed for individuals ages 18 months to 21 years. Internal consistency coefficients reported in the manual ranged from .90--.97 depending on age and test-retest reliability was .96. The Arizona has strong evidence of content, response process, construct, and convergent validity.

Assessment Session. Three reading comprehension and three listening comprehension measures were administered at the second study session (see Table 4). The assessment session usually occurred within one month of the eligibility session; two exceptions were made for one DS and one TD participant who completed the assessment session four months after the eligibility session. Parallel measurement methods, that is measures with similar formats, were selected such that each measurement method was used once in listening comprehension and once in reading comprehension. See Table 1. Assessment order was counterbalanced across participants to control for order effects. Participants were given breaks between tasks as needed to maintain attention and on-task behavior. The assessment session for each participant lasted 75-100 minutes.

Reading comprehension. The Kaufman Ability Battery for Children (KABC; Kaufman & Kaufman, 1983) Reading/Understanding subtest requires a nonverbal response. Test takers act out written directions. The Reading/Understanding subtest is normed for individuals ages 7 to

12. The manual reported mean internal consistency coefficient based on the split-half reliability method based on age of .90 for preschool children and .93 for children ages 5 to 12 years and the test-retest reliability coefficient of .83, .88, and .92 for ages 2;6-4, 5-8, and 9-12;6, respectively. The WRMT-III Passage Comprehension subtest uses cloze procedure. Initial passages are single sentences and passages increase in length across the subtest. Initial passages are accompanied by a picture and pictures are phased out as passages increase in length. Test takers supply the missing word located anywhere in the sentence to complete the meaning of a sentence or paragraph that they read. The manual reported mean internal consistency coefficient based on the split-half reliability method based on age of .90 and the test-retest reliability coefficient of .86 for Pre-Kindergarten-Grade 2, .88 for grades 3-8, and .81 for grades 9-12. Raw scores were calculated on the KABC Reading/Understanding and WRMT-III Passage Comprehension subtests. The Wechsler Individual Achievement Test-III (WIAT-III; Wechsler, 2009) Reading Comprehension subtest uses passage-level text paired with open-ended questions. Test takers read passage-level text and then verbally answer open-ended questions read aloud by the examiner. Test takers' answers were scored according to the criteria provided on the Record Form; answers could be scored as 2-points, 1-point, or 0-points for some questions and scored as 2-point or 0-points on other questions. Four to eight questions were asked per passage. For participants with DS, the entry point was based on their word-level reading grade equivalent based on the WRMT-III Word Identification subtest and for TD participants, the entry point was based on their current grade level. Because WIAT-III Reading Comprehension scores are based on the particular item set administered and the total raw scores from different item sets are not directly comparable, vertically scaled scores (i.e., weighted scores) were used as outlined in the assessment manual. The WIAT-III is normed for individuals ages 4 to 50. The mean internal reliability coefficient reported in the manual was .86 and test-retest reliability was .90.

Listening comprehension. Two subtests from the Woodcock-Johnson IV Test of Oral Language (WJ IV; Schrank et al., 2014), normed for individuals ages 2 to 90 years, were administered. The Understanding Directions subtest requires a nonverbal response. Test takers follow directions presented orally via an audio recording to point to familiar objects with varying characteristics (e.g., size, location) in a picture scene. This subtest has a median reliability of .86 in the 5 to 19 age range and .87 in the adult age range as reported in the manual. The Oral Comprehension subtest uses a cloze procedure. Test takers listen to a short audio-recorded passage and supply the missing word at the end of the sentence. This subtest has a median reliability of .82 in the 5 to 19 age range and .80 in the adult age range. The WIAT-III Listening Comprehension Oral Discourse Comprehension subtest was administered. Test takers listen to audio-recorded passage-level text and then verbally answer open-ended questions read aloud by the examiner. Test takers' answers were scored according to the possible correct answers listed on the Record Form; one point was awarded for each correct answer and zero points for incorrect answers. The mean internal reliability coefficient reported in the manual was .83 and test-retest reliability was .75. The WIAT-III was found to have strong evidence of validity based on content, response process, and internal structure. Raw scores were calculated for all three listening comprehension measures.

Reliability. As the primary coder, the first author scored all measures initially. A graduate student reliability coder with formal training in psychoeducational assessment was trained on the scoring procedures for the dependent measures. She independently scored a random selection ($\geq$ 25%) of the participants' assessment sessions from video and audio recordings; only video recordings with camera angles that allowed for valid assessment scoring were eligible for random selection. Interobserver reliability was estimated using intraclass correlation coefficients (ICCs). ICCs account for differences in scores between coders as well as the variance among participants on the measures of interest. For the dependent measures, the mean ICC value was .99 for the DS group and .93 for the TD group (Hessling, 2020) and thus, the primary coder's scoring was used in the analyses. The ICC values were all excellent for the DS group (.94–1.00) and the values ranged from good to excellent for the TD group (.80–1.00). For both groups, the lowest ICC values were observed for the WIAT-III measures which is not surprising given that the response format is an open-ended verbal response, and thus the rubric requires decisions by the coder which may lead to potential lack of agreement across coders. See Table 7 for ICC values for each measure.

Table 7

ICC Values

	ICC Value	
Measure	DS Group (n = 19)	TD Group (n = 19)
Reading Comprehension		
KABC Reading/Understanding	.99	.99
WRMT-III Passage Comprehension	.99	1.00
WIAT-III Reading Comprehension	.94	.84
Listening Comprehension		
WJ-IV TOL Understanding Directions	1.00	.99
WJ-IV TOL Oral Comprehension	1.00	.96
WIAT-III Listening Comprehension	1.00	.80

Note. ICC = intraclass correlation coefficient; DS = Down syndrome; TD = typically developing; KABC = Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983); WRMT-III = Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); WIAT-III = Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009); WJ-IV TOL = Woodcock Johnson IV Tests of Oral Language (Schrank et al., 2014).

The primary coder and reliability coder double scored all descriptive measures (93% inter-rater agreement) and discrepancies were resolved by consensus before data was double entered for analysis.

Statistical Analyses

The dependent variable for each dependent measure was the participant's raw score with one exception for the WIAT-III Reading Comprehension subtest, as described above. We used raw scores so as to capture incremental differences between participants that would be obscured by using standard scores for individuals with ID (e.g., Kover & Atwood, 2013; Mervis & Klein-Tasman, 2004).

Pearson r correlations were calculated to evaluate the relation between the components of the Simple View of Reading among individuals with DS and TD children. To evaluate between-group differences on the reading comprehension and listening comprehension measures, six independent-samples t-tests were conducted, one for each of the six dependent measures (three reading comprehension subtests and three listening comprehension subtests). A Bonferroni correction was used to control for multiple comparisons with an adjusted alpha level of .008. Cohen's *d* effect sizes were calculated to quantify the magnitude of between-group differences and interpreted in the context of the effect sizes reported in other similar studies (Dunst & Hamby, 2012).

Results

Correlations Among Measures in the DS and TD Groups

Correlational analyses for each group which included word reading and the three reading comprehension and listening comprehension measurement methods are reported in Table 8. For the DS group, moderate to strong significant correlations were observed between word reading and all the reading comprehension and listening comprehension measures, whereas for the TD group, word reading was correlated significantly only with the reading comprehension measures. In evaluating the relation between reading comprehension and listening comprehension, for the DS group, comparable correlations were found across all the reading comprehension and listening comprehension measurement methods. However, for the TD group, only four out of the nine correlations were significant; the extent to which reading comprehension and listening comprehension were correlated varied by measurement method. Reading comprehension as measured with open-ended questions correlated significantly with all three listening comprehension measures.

Table 8

Relation among word reading, reading comprehension, and listening comprehension

		1	2	3	4	5	6	7
Down Syndrome Group								
1. Word Reading	WRMT-III Word ID	1	.78**	.73**	.70**	.67*	.59**	.54*
2. RC Nonverbal	KABC Reading/ Understanding		1	.85**	.77**	.79**	.71**	.65**
3. RC Cloze Procedure	WRMT-III Passage Comprehension			1	.81**	.75**	.70**	.68**
4. RC Open-Ended	WIAT-III Reading Comprehension				1	.64**	.68**	.69**
5. LC Nonverbal	WJ-IV TOL Understanding Directions					1	.90**	.78**
6. LC Cloze Procedure	WJ-IV TOL Oral Comprehension						1	.88**
7. LC Open-Ended	WIAT-III Listening Comprehension							1
Typically Developing Group								
1. Word Reading	WRMT-III Word ID	1	.85**	.85**	.77**	.39	.43	.42
2. RC Nonverbal	KABC Reading/ Understanding		1	.78**	.88**	.41	.43	.38
3. RC Cloze Procedure	WRMT-III Passage Comprehension			1	.81*	.41	.58**	.44
4. RC Open-Ended	WIAT-III Reading Comprehension				1	.55*	.61**	.57*
5. LC Nonverbal	WJ-IV TOL Understanding Directions					1	.62**	.70**
6. LC Cloze Procedure	WJ-IV TOL Oral Comprehension						1	.66**
7. LC Open-Ended	WIAT-III Listening Comprehension							1

Note. RC = Reading Comprehension; LC = Listening Comprehension; KABC = Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983); WRMT-III = Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); WIAT-III = Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009); WJ IV = Woodcock-Johnson IV (Schrank et al., 2014); *$p < .05$, **$p < .01$.

Group Differences in Reading Comprehension

Independent-samples t-tests with Bonferroni adjusted alpha levels of .008 were conducted to compare the DS and TD groups on reading comprehension. Despite that participants were matched on word reading raw scores ($t(38) = .21$, $p = .83$), there was a significant between-group difference for two out of the three reading comprehension measures (see Table 9). The reading comprehension mean raw scores for the DS group were lower than for the TD group, with large effect sizes (*d*s = 1.17, 1.57). Figure 1 shows the results for each group on each respective reading comprehension measure. The violin plots illustrate kernel density estimation, that is, wider sections of the violin plot represent a higher probability that members of the population will take on the given value; the skinnier sections represent a lower probability.

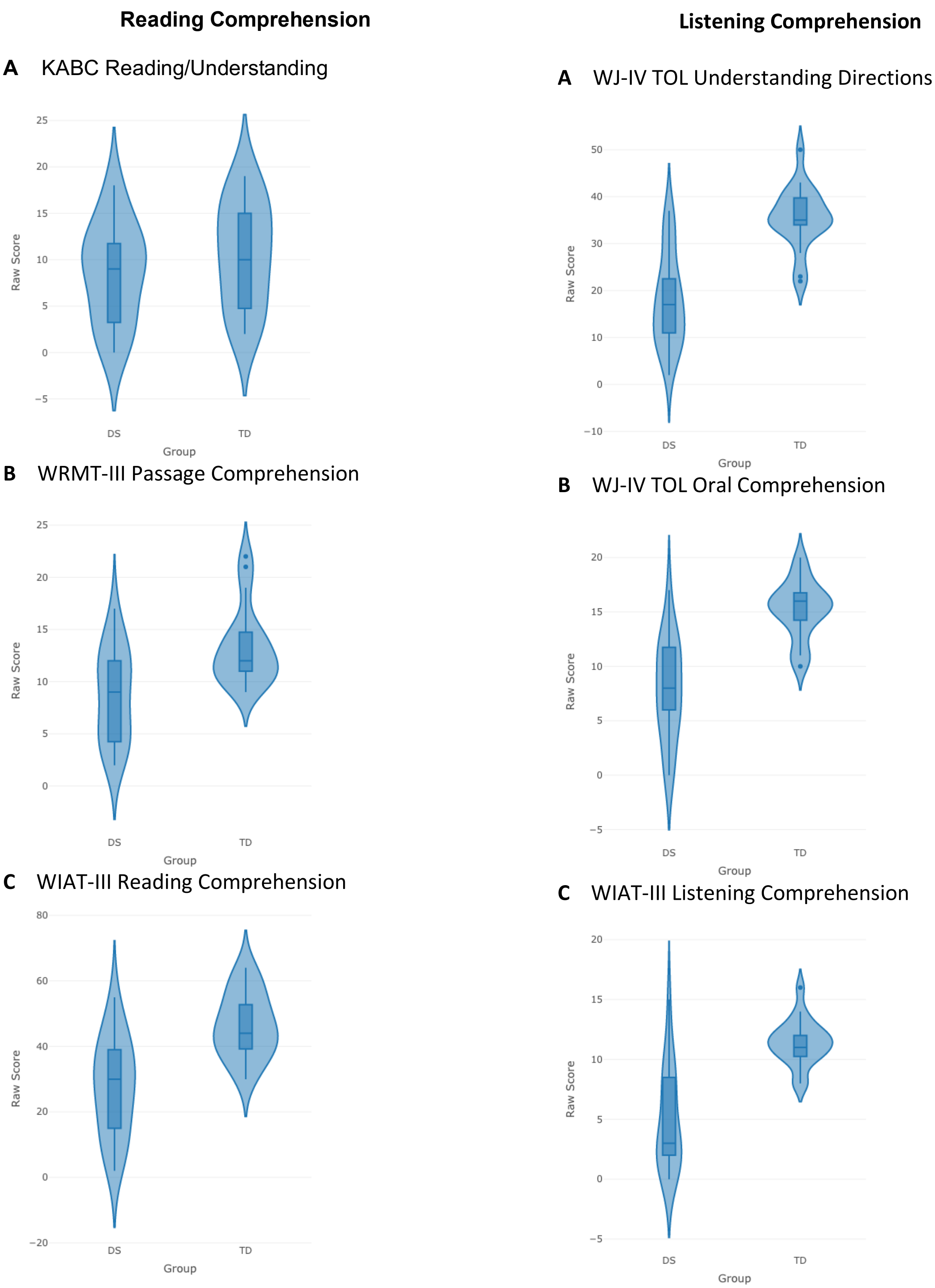

Figure 1. Violin plots of raw scores on each reading comprehension and listening comprehension measure by group. DS = Down syndrome, TD = typical development. The data are presented as box plots, indicating the median (horizontal line) and quartiles with whiskers reflecting 1.5 times the interquartile range. Violin plots are shown to visualize the distribution of data and its probability density (the width of the colored area represents the proportion of the data located here) for each measure and each group.

Group Differences in Listening Comprehension

Independent-samples t-tests with Bonferroni adjusted alpha levels of .008 were conducted to compare the DS and TD groups on listening comprehension. There was a significant between-group difference for all listening comprehension measures (see Table 9) The three listening comprehension mean raw scores for the DS group were lower than the TD group, with large effect sizes (*d*s = 2.03 – 2.17). Notably, the between-group effect sizes were larger for listening comprehension than reading comprehension. Figure 1 shows the results for each group on each respective listening comprehension measure. For the DS group, a greater proportion of the data was located near the floor level (i.e., at zero) only on the WIAT Listening Comprehension measure.

Table 9

Participant Reading Comprehension and Listening Comprehension Raw Scores

	DS Group (n = 19)			TD Group (n = 19)					
Measure	Mean	SD	Range	Mean	SD	Range	*t*	*p*	*d*
Reading Comprehension									
KABC Reading/Understanding	8.58	5.32	0-18	10.58	5.64	2-19	1.125	.27	0.38
WRMT-III Passage Comprehension	8.68	4.41	2-17	13.32	3.73	9-22	3.50	.00	1.17
WIAT-III Reading Comprehension*	27.42	14.67	2-55	46.37	9.71	30-64	4.70	.00	1.57
Listening Comprehension									
WJ-IV TOL Understanding Directions	17.74	9.89	2-37	35.47	6.53	22-50	6.35	.00	2.17
WJ-IV TOL Oral Comprehension	7.84	4.62	0-17	15.26	2.62	10-20	6.09	.00	2.03
WIAT-III Listening Comprehension	4.95	4.13	0-15	11.42	1.90	8-16	6.21	.00	2.07

Note. *vertically scaled scores (not raw scores) reported for this measure, due to administration rules; DS = Down syndrome; TD = typically developing; KABC = Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983); WRMT-III = Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); WIAT-III = Wechsler Individual Achievement Test–Third Edition (Wechsler, 2009); WJ IV = Woodcock-Johnson IV (Schrank et al., 2014).

Discussion

Relation Between Word Reading, Listening Comprehension, and Reading Comprehension

In this study, we evaluated the relation of word reading, reading comprehension, and listening comprehension within a group of individuals with DS and within a group of TD children who were matched on word reading to the DS group. The study design allowed us to consider these relations in two groups with similar word reading skills.

Word reading and reading comprehension, word reading and listening comprehension. Within the DS group, word reading was strongly correlated with reading comprehension (*r*s = .69 - .78) as well as listening comprehension (*r*s = .77 - .85). However within the TD group, although word reading was strongly correlated with reading comprehension (*r*s = .77 - .85), it was not correlated with listening comprehension. Although significant correlations have not been observed between word reading and listening comprehension in past research with DS, our findings are consistent with previous studies that evaluated the relation between word reading and reading comprehension. Similar to our findings, strong correlations between word reading and reading comprehension have been observed in DS (Boudreau, 2002 [*r* = .66]; Laws et al, 2016 [*r* = .65]) and TD (e.g., Catts et al., 2015 [*r* = .85]). As such, for both groups we can hypothesize that growth in word reading would lead to changes in reading comprehension. Due to the multiplicative effect in the Simple View, we hypothesize that improved listening comprehension may lead to changes in reading comprehension above and beyond the changes from growth in word reading for the DS group. Thus, these findings provide a strong rationale for an investigation of the unique correlations and causal effects among word reading, listening comprehension, and reading comprehension.

Reading comprehension and listening comprehension. We found that reading comprehension and listening comprehension were moderately to highly correlated within participant groups. Shared variance across measurement methods for the DS group ranged from 41 – 62% and for the TD group 30 – 37%. For the DS group reading comprehension was strongly correlated with listening comprehension across all measurement methods, whereas for the TD group, moderate correlations were found for only four out of the nine comparisons. As such, the impact of listening comprehension on reading comprehension may be stronger and more consistently observed, regardless of text format or measurement method, in DS than TD. Although Laws and colleagues (2016) reached a similar conclusion, Roch and Levorato (2009) observed moderate correlations between reading comprehension and listening comprehension for both groups, DS (r = .41) and TD (r = .52). Although the percentage of shared variance was greater in the TD group (r^2 = 27%) compared with the DS group (r^2 = 17%), based on hierarchical regression analyses, they found that in the DS group, only listening comprehension and not word reading predicted reading comprehension (R^2 = .192, ß = .438, p < 0.001). In the context of the Simple View, these findings further suggest that improving listening comprehension would have a greater impact on reading comprehension in the DS group compared with the TD group. Taken together, these findings support the Simple View of Reading in DS in that word reading and listening comprehension are related to reading comprehension. The DS-TD group comparisons resulted in some differences in the magnitude of the relations between the Simple View components. Thus, as is suggested in other theoretical models, such as Scarborough's reading rope, there are likely other underlying cognitive and linguistic skills that contribute to reading comprehension that we did not account for (Scarborough, 2001). Further exploration into the extent to which underlying skills (e.g., cognition, memory, oral language) beyond those included in the Simple View contribute to reading comprehension is necessary to advance an understanding of reading development in DS.

Group Differences in Reading Comprehension and Listening Comprehension

Previous research comparing reading comprehension and listening comprehension in individuals with DS and TD children has been limited. In past studies, researchers evaluated reading comprehension and listening comprehension only with sentence-level text and in younger readers who may not have had foundational reading skills to score above floor levels (Laws et al., 2016). Researchers have also previously evaluated these constructs in Italian-speaking individuals with DS (Roch & Levorato, 2009; Roch et al., 2011). When evaluating between-group differences, we found that the DS group, despite being matched on word reading, demonstrated lower reading comprehension scores compared with the TD group on two out of the three measurement methods (*d*s = 1.17, 1.57). Laws et al (2016) reported an effect size of a similar magnitude (*d* = 1.69) to the larger of the two effect sizes reported in our study (*d* = 1.57) whereas Nash and Heath (2011) reported a larger effect size (*d* = 2.34) in their sample of older individuals with DS (11- to 19-year-olds). However, it is important to note that Nash and Heath's TD comparison group was older (8- to 9-year-olds) compared with our TD group (6- to 8-year-olds). Our smaller between-group reading comprehension effect sizes may be explained by our attempts to minimize floor effects. For example, greater between-group differences would be observed if many participants in the DS group scored at floor levels and the TD group participants attained age-expected scores, this would be especially apparent in Nash and Heath's older 8- to 9-year-old TD group. Overall, although our DS group presented with at least foundational word reading skills, they do not have reading comprehension comparable to word reading-matched peers.

Additionally, our DS group demonstrated lower listening comprehension scores compared to the TD group on all three measurement methods (*d*s = 2.03 - 2.17). Laws and colleagues (2016) observed a group difference with effect sizes of a similar magnitude for listening comprehension (*d* = 2.79) relative to sentence-level text. The DS group's weaknesses in listening comprehension, regardless of measurement method, likely presents a barrier to reading comprehension development and is a source for discrepant reading comprehension when compared to TD children. Given the between-group difference in reading comprehension (*d* = 1.17, 1.57), we would expect a between-group difference in listening comprehension based on the Simple View. However, it is surprising that we found a much larger discrepancy between the TD and DS groups for listening comprehension (*d*s = 2.03 – 2.17). than for reading comprehension (*d* = 1.17, 1.57) as evidenced by greater *t*-statistics and effect sizes (though nearly all independent sample *t*-tests were significant and all the effect sizes were large). This finding is consistent with Laws et al (2016) on their sentence-level comprehension measures and again points to the importance of considering listening comprehension as a potentially malleable factor to promote reading comprehension.

Consideration of the pattern of strengths and weaknesses with the DS phenotype also may explain the greater between-group difference in listening comprehension compared with reading comprehension. More so than in reading comprehension, successful listening comprehension may tap certain underlying skills such as oral language, short-term memory, and working memory, all of which are distinct areas of need in the DS phenotype. Likewise, individuals with DS may optimize their relative strength in visual processing in reading comprehension tasks in which the text provides visual support compared with listening comprehension tasks which relies on information presented orally. Future research should investigate the role of these distinct phenotypic characteristics in listening comprehension and reading comprehension in DS.

Clinical Implications

Several implications for speech-language pathologists and special educators emerge from our findings. To optimize treatment effects, it is imperative to consider the DS phenotype. The results of the present study provide further support for utilizing a DS-specific approach which has successfully been implemented in reading interventions for younger children with DS (e.g., Lemons et al., 2015). Because children with DS often present with language deficits, it is critical that intervention not only focus on word reading but includes an explicit focus on listening comprehension as well. Speech-language pathologists and special educators may be inclined to build upon a relative strength in sight word-recognition for a child with DS and focus solely on this dimension in the beginning stages of intervention, thus disregarding or minimizing other literacy skills such as decoding and listening comprehension This inclination likely is based on the ideas that comprehending text is (a) predicated on proficient and robust word-recognition and (b) more difficult to teach compared with increasing sight word vocabulary. However, the relation between word reading and listening comprehension within the Simple View of Reading and our findings suggest that targeting *both* word reading and listening comprehension is necessary to facilitate proficient reading comprehension.

Rather than simply building a robust word recognition vocabulary and engaging children in learning to read with texts limited to those words, interventionists must challenge children with DS to simultaneously develop reading comprehension. In particular, incorporating appropriate and meaningful texts, specifically texts that contain a plot and thus necessitate comprehension, is vital for children with DS (Morgan & Moni, 2008). Even though meaningful texts may include words that a child cannot yet read, encountering these texts presents an opportunity to engage in listening comprehension in which the child comprehends the text that is read-aloud by the interventionist. The use of specifically-designed teaching strategies and resources such as continuing to read to children while simultaneously teaching them to read using meaningful texts may be critical for children with DS to develop text comprehension. Thus, interventionists need to support reading comprehension, word-level reading, and perhaps most importantly, listening comprehension in treatment to optimize reading outcomes in DS.

Limitations and Future Directions

Limitations of the present study provide future directions for this line of inquiry. First, participants were not matched on socioeconomic status (SES), although SES-matching should be considered in future studies. Second, six individuals with DS who were consented did not meet the study word-level reading eligibility criteria. To fully understand the literacy skills of the DS population, longitudinal studies are needed that capture the course of literacy development. Our study was not designed to capture the range of reading skills, yet understanding the listening comprehension skills of students with DS who are not yet readers may be important to designing intervention. Further, longitudinal studies can advance understanding of what differentiates individuals with DS who develop functional reading skills from those who do not.

Third, the study results are specific to a particular subset of individuals with DS within a specific developmental period and the degree to which these results can be generalized for a broader and more representative sample in individuals with DS is unknown. Thus, it is essential to interpret these findings in reference to development, which elucidates the need to evaluate the discrepancy in these domains—word-level reading, listening comprehension, and reading comprehension—over time and across multiple developmental periods. In addition, future research should evaluate listening comprehension as a potentially malleable factor to improve reading comprehension outcomes for individuals with DS.

Summary

Our findings demonstrate that readers with DS have a specific phenotypic pattern of strengths and weaknesses in word reading, listening comprehension, and reading comprehension, providing additional support for Lemons and colleagues' (2015) argument for intervention protocols that consider phenotypic characteristics. Intervention protocols for TD readers, for example, may not focus sufficiently on listening comprehension for DS readers. In general, the assumption is that children with TD have a robust capacity for change, whereas children with DS do not and thus their outcomes are more reliant on the specific interventions that they receive (e.g., Grieco et al., 2015; Lemons et al, 2015). As such, for individuals with DS, intervention must focus not only on word reading, but also on listening comprehension as it relates directly to reading comprehension. As a word of caution, it is important to consider the distinction between supporting oral language or language comprehension versus listening comprehension (e.g., LARRC, 2017). Oral language comprehension involves redundancy in which speakers revise the message presented when a listener does not comprehend, whereas this opportunity to repair communication breakdowns is not present in text-based listening comprehension. Listening comprehension intervention is text-based and involves supporting students to construct meaning from read-aloud written text (Gough & Tunmer, 1986) which distinguishes it from receptive language or language comprehension intervention involving vocabulary or grammar (e.g., LAARC, 2017), for example.

Additionally, our results reveal measure-specific information. First, given the small effect size for the nonverbal response reading comprehension measure (KABC Reading/ Understanding), this measure may reduce cognitive-linguistic demands given that participants have to act out their response. As such, this measure may not provide the best indication of an individuals' reading comprehension in academic settings and thus it may underestimate the challenges an individual with DS has with reading comprehension. Second, compared to our findings (d = 1.17, 1.57), Nash and Heath (2011) observed a larger group difference in reading comprehension (d = 2.34) as measured by the Neale Analysis of Reading Ability – Revised (NARA II, Neale, 1997). Between-group effect sizes may be partially dependent on the measurement method. For instance, a larger between-group difference may be observed for a measure that involves reading a short story followed by answering comprehension questions (such as the NARA) as compared with the reading comprehension measurement methods included in the current study. Because all reading comprehension measures do not tap the same aspects of comprehension, educators and researchers need to choose reading comprehension measures wisely, based on what they specifically want to know (e.g., Cutting & Scarborough, 2006; Elleman & Compton, 2017). This consideration may be especially important to better understand how various underlying language skills relate to reading comprehension in DS.

REFERENCES

Allor, J. H., Gifford, D. B., Al Otaiba, S., Miller, S. J., & Cheatham, J. P. (2013). Teaching students with intellectual disability to integrate reading skills: Effects of text and text-based lessons. *Remedial and Special Education, 34*(6), 346–356. https://doi.org/10.1177/0741932513494020

Allor, J.H., Mathes, P.G., Roberts, J.K., Cheatham, J.P., & Champlin, T.M. (2010). Comprehensive reading instruction for students with intellectual disabilities: Findings from the first three years of a longitudinal study. *Psychology in the Schools, 47*(5), 445-466. https://doi.org/10.1002/pits.20482

American Speech-Language-Hearing Association (n.d.). *Childhood Hearing Screening*. Retrieved May 1, 2021, from https://www.asha.org/PRPSpecificTopic.aspx?folderid=8589935406&section=Key_Issues.

Avalos, M. A., Plasencia, A., Chavez, C., & Rascón, J. (2007). Modified guided reading: Gateway to English as a second language and literacy learning. *The Reading Teacher*, *61*(4), 318-329.

Boudreau, D. (2002). Literacy skills in children and adolescents with Down syndrome. *Reading and Writing, 15*(5), 497-525.

Carrow-Woolfolk, E. (2014). *Test for Auditory Comprehension of Language-Fourth Edition (TACL-4)*. Pro-Ed.

Catts, H. W., Herrera, S., Nielsen, D. C., & Bridges, M. S. (2015). Early prediction of reading comprehension within the simple view framework. Reading and Writing, 28(9), 1407–1425. http://dx.doi:10.1007/s11145-015-9576-x

Catts, H. W., & Kamhi, A. G. (2017). Prologue: Reading comprehension is not a single ability. *Language, Speech, and Hearing Services in Schools*, *48*(2), 73–76. https://doi.org/10.1044/2017_LSHSS-16-0033

Chapman, R. S., & Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, *6*(2), 84-95. https://doi.org/10.1002/1098-2779(2000)6:2<84::AID-MRDD2>3.0.CO;2-P

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*(3), 277-299. https://doi.org/10.1207/s1532799xssr1003_5

Dunst, C. J., & Hamby, D. W. (2012). Guide for calculating and interpreting effect sizes and confidence intervals in intellectual and developmental disability research studies. *Journal of Intellectual & Developmental Disability*, *37*(2), 89–99. https://doi.org/10.3109/13668250.2012.673575

Elleman, A. M., & Compton, D. L. (2017). Beyond comprehension strategy instruction: What's next? *Language, Speech, and Hearing Services in Schools*, *48*(2), 84-91. https://doi.org/10.1044/2017_LSHSS-16-0036

Fitzgerald, J. (2003). Multilingual reading theory. *Reading Research Quarterly*, *38*(1), 118-122.

Fuchs, D., Hendricks, E., Walsh, M. E., Fuchs, L. S., Gilbert, J. K., Zhang Tracy, W., Patton, S., Davis-Perkins, N., Kim, W., Elleman, A., & Peng, P. (2018). Evaluating a multidimensional reading comprehension program and reconsidering the lowly reputation of tests of near-transfer. *Learning Disabilities Research & Practice*, *33*(1), 11-23. https://doi.org/10.1111/ldrp.12162

Fudala, J., & Stegali, S. (2017). *Arizona Articulation and Proficiency Scale–Fourth Revision (Arizona–4)*. Western Psychological Services.

Gough, P. & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6-10. https://doi.org/10.1177/074193258600700104

Grieco, J., Pulsifer, M., Seligsohn, K., Skotko, B., & Schwartz, A. (2015, June). Down syndrome: Cognitive and behavioral functioning across the lifespan. *American Journal of Medical*

Genetics Part C: Seminars in Medical Genetics, 169(2), 135-149. http://dx.doi.org/10.1002/ajmg.c.31439

Hessling, A. (2020). *The not so simple view of reading in Down syndrome*. [Doctoral dissertation, Vanderbilt University]. ProQuest Dissertations Publishing.

Invernizzi, M., Meier, J.D., Swank, K., & Juel, C. (1997). *Phonological Awareness Literacy Screening (PALS).* Charlottesville, VA: University Printing Services.

Kamhi, A.G., & Catts, H.W. (1999). Defining reading disabilities. In H.W. Catts & A.G. Kamhi (Eds.), *Language and reading disabilities* (pp. 50-72). Boston: Allyn & Bacon.

Kaufman, A. S. (2004). *Kaufman Brief Intelligence Test–Second Edition (KBIT-2).* Pearson

Kaufman, A.S. & Kaufman, N.L. (1983). *Kaufman Assessment Battery for Children*. American Guidance Service.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kover, S. T., & Atwood, A. K. (2013). Establishing equivalence: Methodological progress in group-matching design and analysis. *American journal on intellectual and developmental disabilities*, *118*(1), 3-15.

Language and Reading Research Consortium. (2017). Oral language and listening comprehension: Same or different constructs?. *Journal of Speech, Language, and Hearing Research*, *60*(5), 1273-1284. http://dx.doi.org/10.1044/2017_JSLHR-L-16-0039

Laws, G., & Gunn, D. (2002). Relationships between reading, phonological skills and language development in individuals with Down syndrome: A five-year follow-up study. *Reading and Writing, 15*(5), 527-548.

Laws, G., Brown, H., & Main, E. (2016). Reading comprehension in children with Down syndrome. *Reading and Writing*, *29*(1), 21-45. https://doi.org/10.1007/s11145-015-9578-8

Lemons, C. J., King, S. A., Davidson, K. A., Puranik, C. S., Al Otaiba, S., Fulmer, D., Mrachko, A.A., Partanen, J., & Fidler, D. J. (2017). Developing an early reading intervention aligned with the Down syndrome behavioral phenotype. *Focus on Autism and Other Developmental Disabilities*, *32*(1), 176-187. https://doi.org/10.1177/1088357615618941

Lemons, C. J., King, S. A., Davidson, K. A., Puranik, C. S., Fulmer, D., Mrachko, A. A., Partanen, J., Otaiba, S. A., & Fidler, D. J. (2015). Adapting phonological awareness interventions for children with Down syndrome based on the behavioral phenotype: A promising approach?. *Intellectual and Developmental Disabilities, 53*(4), 271-288. https://doi.org/10.1352/1934-9556-53.4.271

Martin, N., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test, Fourth Edition (EOWPVT-4)*. Academic Therapy Publications.

Martin, N., & Brownell, R. (2011). *Receptive One-Word Picture Vocabulary Test, Fourth Edition (ROWPVT-4)*. Academic Therapy Publications.

Megherbi, H., Seigneuric, A., & Ehrlich, M.F., 2006, Reading comprehension in French 1st and 2nd grade children: Contribution of decoding and language comprehension. *European Journal of Psychology and Education, 21*(2), 135-147. https://doi.org/10.1007/BF03173573

Mervis, C. B., & Klein-Tasman, B. P. (2004). Methodological issues in group-matching designs: α levels for control variable comparisons and measurement characteristics of control and target variables. *Journal of Autism and Developmental Disorders*, *34*(1), 7-17.

Mervis, C. B., & Robinson, B. F. (2003). Methodological issues in cross-group comparisons of language and cognitive development. In Y.Levy & J.Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 233-258). Erlbaum.

Nash, H., & Heath, J. (2011). The role of vocabulary, working memory and inference making ability in reading comprehension in Down syndrome. *Research in Developmental Disabilities*, *32*(5), 1782-1791. https://doi.org/10.1016/j.ridd.2011.03.007

Neale, M. D. (1997). Neale Analysis of Reading Ability – Revised (2nd ed.). NFER-Nelson.

Roch, M., & Levorato, M. C. (2009). Simple view of reading in Down's syndrome: The role of listening comprehension and reading skills. *International Journal of Language & Communication Disorders*, *44*(2), 206-223. https://doi.org/10.1080/13682820802012061

Roch, M., Florit, E., & Levorato, C. (2011). A follow-up study on reading comprehension in Down syndrome: the role of reading skills and listening comprehension. *International Journal of Language & Communication Disorders*, *46*(2), 231-242. https://doi.org/10.3109/13682822.2010.487882

Roizen, N. J., Wolters, C., Nicol, T., & Blondis, T. A. (1993). Hearing loss in children with Down syndrome. *The Journal of Pediatrics*, *123*(1), S9-S12. https://doi.org/10.1016/S0022-3476(05)81588-4

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis) abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy* (pp. 97-110). New York: Guilford Press.

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Oral Language*. Riverside Insights.

Semel, E., Wiig, E., & Secord, W. (2000). *Clinical Evaluation of Language Fundamentals—3UK*. Pearson.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Rand Corporation.

Wechsler, D. (2009). *Wechsler Individual Achievement Test–Third Edition (WISC-3).* Pearson.

Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests-Third edition (WRMT-III)*. Pearson.