Title

Reading enhancements for students with autism spectrum disorder: A matched randomized pilot experimental study

Authors

Solis, Michael
Reutebuch, Colleen
Vaughn, Sharon
et al.

Publication Date

2022-05-01

DOI

10.1016/j.rasd.2022.101937

Peer reviewed

Reading Enhancements for Students with Autism Spectrum Disorder: A Matched Randomized Pilot Experimental Study

Michael Solis

University of California Riverside, corresponding author, msolis@ucr.edu

1207 Sproul Hall, Riverside, CA 92521

Colleen Reutebuch, Sharon Vaughn

The University of Texas at Austin

1912 Speedway, Austin, TX 78712

Zaira Jimenez

University of California Riverside

1207 Sproul Hall, Riverside, CA 92521

Author Note

Michael Solis, Graduate School of Education, University of California Riverside; Colleen K. Reutebuch, Meadows Center for Preventing Educational Risk, The University of Texas Austin; Sharon Vaughn, Department of Special Education, The University of Texas Austin; Zaira Jimenez, University of California Riverside

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A160299 to the University of California Riverside. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

Correspondence concerning this article should be addressed to Michael Solis, 1207 Sproul Hall, University of California Riverside, Riverside, CA 92521. E-mail: msolis@ucr.edu

Abstract

Background: For students with ASD, improving reading comprehension is critical, increasing both their opportunities for successful postsecondary experiences including attending college and obtaining meaningful employment.

Method: We conducted a matched randomized pilot experimental study for students with ASD in grades 3 - 8 (N=28). An independent researcher matched participants according to symptom severity and reading fluency and then randomly assigned a member of each pair to treatment or comparison condition. Participants in treatment condition were provided 1:1 instruction for 23 – 30 sessions (M = 27) of 30 minutes each four to five days per week. Instructional components included (a) vocabulary instruction; (b) fluency with text, and (c) reading comprehension.

Results: The intervention was associated with significant gains in reading comprehension [WJ-PC scores of 22.62 (*se*=8.19, *df*=3, *p*=.070) for students at GARS severity level 1; *p* value < 0.10], and with significant gains in vocabulary scores of 10.19 (*se*=2.78, *df*=3, *p*=.035) and 5.46 (*se*=1.60, *df*=3, *p*=.042) for students rated at GARS severity levels 1 and 2. Significant effects were not detected for scores on researcher-developed measure of reading comprehension or a standardized measure of reading and fluency (TOSREC).

Conclusions: Although growth is modest and appears to be limited to participants who are in the mid to higher range of the autism spectrum, considering the need for evidence-based practices we interpret the findings as having implications for instruction.

Reading Enhancements for Students with Autism Spectrum Disorder: A Matched Randomized Pilot Experimental Study

The prevalence rates of autism spectrum disorder (ASD) across the world were recently estimated to be approximately one of every 132 persons (Baxter et al., 2015). The global rates of ASD continue to rise based on data from developed and developing countries (Onaolapo & Onaolapo, 2017). In comparison to other disabilities, the global prevalence of ASD in children is greater than conduct disorder and attention deficit hyperactive disorder (ADHD) combined (Baxter et al., 2015).

In the United States the prevalence rates of ASD for children continue to increase dramatically (1 in 88 children in 2008, 1 in 68 children in 2014, and 1 in 59 children in 2018; Center for Disease Control and Prevention, n.d.). The increased prevalence of children with ASD require the need for empirically validated interventions, specifically for the educationally high priority area of reading development. In light of the landmark Supreme Court case providing clarification regarding the interpretation of the Free and Appropriate Public Education (FAPE) clause (Yell & Bateman, 2017), school districts are now required to go above and beyond the previous held "de minimis" standard. Thus, educators need access to more specialized, evidence-based instruction as part of their specialized instructional supports required to improve academic outcomes.

For students with ASD, improving reading comprehension is critical, increasing both their opportunities for successful postsecondary experiences including attending college and obtaining meaningful employment. Higher levels of reading comprehension are associated with greater gains in other academic areas, higher levels of employment, increased independence, and overall improved quality of life (Lyon, 1998). Improved academic performance may be

just as important as social skills training for individuals with ASD (Grandin et al., 2004). This study contributes to the small yet growing body of literature investigating reading interventions designed to improve reading performance for students with ASD.

Theoretical Framework and Empirical Underpinnings

Through a multi-year iterative development process, we developed a multicomponent reading intervention derived from several theoretical and empirical lines of research: (a) theoretical underpinnings of reading instruction derived from the Simple View of Reading (SVR; Gough & Tunmer, 1986), (b) the linkage to learning challenges of many students with ASD as informed by the Weak Central Coherence cognitive processing theory (WCC; Happe & Frith, 2006; Koegel, et al., 1999; Quill, 2000), and (c) empirical findings from reader profile studies of students with ASD (Huemer & Mann, 2010; Lucas & Norbury 2014; McIntyre et al., 2017; Randi et al., 2010; Ricketts et al., 2013; Solari et al., 2017; Solari et al., 2019).

Theoretical framework. The Simple View of Reading (SVR) posits that reading comprehension is the product of word recognition and linguistic comprehension (Gough & Tunmer, 1986). Weaknesses in one or both lead to poor reading comprehension (Catts et al., 2003). There are several studies that replicate support for SVR with samples of students with ASD (Brown et al., 2013; Huemer & Mann, 2010; Lucas & Norbury 2014; McIntyre et al., 2017; Randi et al., 2010; Ricketts, 2011; Sorenson Duncan et al., 2021). For example, a meta-analysis by Brown et al., (2013) comparing neurotypical students to students with ASD reported the two strongest predictors of reading comprehension to be decoding (55% of the variance) and semantic knowledge (57% of the variance). Brown et al., (2013) concluded that taking into account language ability is essential to understanding the reading performance of students with ASD. These findings were further extended by Sorenson Duncan et al., (2021) by taking into

account the associations between several subcomponents of language (vocabulary, morphology, syntax, pragmatics, and listening comprehension) and the strength of these associations with reading comprehension. The findings from Sorensen et al., (2021) further support the importance of taking into account oral language skills as an essential consideration with reading comprehension for students with ASD.

A second theoretical model often applied to students with ASD is the cognitive processing theory of weak central coherence (WCC) (Happe & Frith, 2006; Koegel et al., 1999; Quill, 2000). WCC posits that many students with ASD have difficulty summarizing salient points, understanding main ideas, and/or overly focus on extraneous details (Happe, 2005; Williamson et al., 2009). The WCC processing theory explains ASD in terms of students' difficulties in recognizing big ideas and/or relating big ideas with details. The SVR integrated with the WCC provided an initial heuristic for designing instructional practices responding to the needs of students with ASD. The WCC cognitive processing theory integrates with both sides of the SVR in that it helps to inform decoding and linguistic development.

Empirical underpinnings – reader profile studies. While studies from 30 years ago generally supported the idea that students with ASD demonstrated profiles of high decoding and low comprehension (e.g., Frith & Snowling, 1983), more recent studies report greater heterogeneity in students' performance on word reading and comprehension measures while also accounting for additional factors such as language and other components of reading (i.e., fluency) (Solari et al., 2017). Larger sample sizes and longitudinal studies of linguistic profiles afforded opportunities for more sophisticated analyses such as the use of latent variables and model fit indices to inform interactions and make predictions (Huemer & Mann, 2010; Lucas & Norbury 2014; McIntyre et al., 2017; Randi et al., 2010; Ricketts, 2011; Solari et al., 2017; Solari

et al., 2019). The empirical findings from these studies informed the development of this multicomponent intervention with consideration of the following: (a) the heterogeneity of students with ASD in both the word reading and linguistic comprehension components of the SVR (Huemer & Mann, 2010; Mcyntyre et al., 2017; Solari et al., 2019), (b) the importance of taking into account language development (Lucas & Norbury, 2014; Mcyntyre et al., 2017; Norbury & Nation, 2011; Ricketts et al., 2013), and (c) the evidence suggesting that cognitive phenotype and social communicative factors (i.e., GARS scores), and reading fluency scores inform predictions of reading comprehension for students with ASD (Mcyntyre et al., 2017; Solari et al., 2017).

Reader profile studies of students with ASD have validated the presence of four distinct reader profiles that align with the heuristic of the SVR (Huemer & Mann, 2010; Mcyntyre et al., 2017; Solari et al., 2019), while also highlighting the diversity of performance as considerations for conceptualizing instructional needs. Further, ASD symptomology is an important factor to consider in relation to the role of language development and reading outcomes (Huemer & Mann, 2010; Mcyntyre et al., 2017; Solari et al., 2019). Ricketts et al., (2013) conducted regression analysis with a sample of older students with ASD which indicated that oral language explained unique variance in reading comprehension (Ricketts et al., 2013). Other reader profile studies have reported that language impairment of children with ASD is associated with poor performance in reading outcomes (Lucas & Norbury, 2014; Mcyntyre et al., 2017; Norbury & Nation, 2011). Therefore, it is necessary to account for the role of language within the context of reading comprehension (Lucas & Norbury, 2014; Mcyntyre et al., 2017; Norbury & Nation, 2011; Ricketts et al., 2013). It is also necessary to account for the role of ASD symptomology for other areas of reading including reading fluency (Solari et al., 2017; Solari et al., 2019). Solari et

al., (2017) reported reading fluency as an important consideration for reading comprehension for students with ASD based greater variation being observed compared to other components of reading (Solari et al., 2017). When controlling for both decoding and language, the structural equation models reported by Solari et al., (2019) suggest that reading fluency is a significant predictor of reading comprehension for students with ASD. Therefore, controlling for reading fluency affords an opportunity to isolate on the dependent variables associated with reading comprehension (i.e., vocabulary).

Reading Intervention Research of ASD

Systematic reviews of reading interventions for students with ASD conducted over the last 15 years have contributed to our understanding of the efficacy of particular instructional approaches in reading for these students (Bailey & Ariciuli, 2020; Brown et al., 2013; Chiang & Lin, 2007; El Zein, et al., 2014; Finnegan & Mazin, 2016; Knight et al., 2013; Senokossoff, 2016; Whalon et al., 2009). The majority studies represented across these systematic reviews were implemented with researchers providing the intervention. Approximately, 75% of the studies provided 1:1 instruction for the reading intervention. The literature on reading interventions for students with ASD remains underdeveloped with an over reliance on single-case design studies. However, findings from these studies do provide an initial empirical base for utilizing modeling, guided practice, and independent practice as a mechanism to support improvements in the following areas: vocabulary and graphic organizers (Dugan et al., 1995; Grindle et al., 2013; Kamps et al., 1995; Knight et al., 2015; Williamson et al., 2015), fluency with text (Barnes & Rehfeldt, 2013; Kamps et al., 1994; Kamps et al., 1989; Reisner et al., 2014), sentence comprehension, multi-paragraph comprehension, and questioning strategies (Asberg et al., 2010; Bailey et al., 2017; Bethune & Wood, 2013; Ganz & Flores, 2009;

Reutebuch et al., 2015; Roux et al., 2015; Roux, Dion, & Barrette, 2015; Turner et al., 2017; Whalon & Hanline, 2008). See Table 1 for a summary of studies of reading interventions for students with ASD in the middle grades. These studies provide empirical support for the instructional components of vocabulary, fluency and reading comprehension

It is also important to note the small number of pre-post group design studies included in these systematic reviews and the lack of student samples with older students from upper elementary and middle school grades. Three of these studies focused on younger students (average age 9 yrs.) (Bailey & Ariciuli, 2017; Roux et al., 2015; Roux, Dion, & Barrette, 2015) with only one study utilizing a sample of older students (age 13 years) (Turner et al., 2017). There are also methodological limitations to these studies. For example, Turner et al., (2017) employed a pre-post design with no comparison group which eliminates experimental control to isolate the impact of the intervention compared to maturation. Nevertheless, the remaining studies with multiple conditions did report improvements favoring the intervention conditions compared to business as usual conditions. This pilot study for students with ASD in the middle grades contributes to the literature by testing a multicomponent intervention that, while similar in some facets, is unique in terms of the makeup of the instructional components compared to previous approaches to intervention.

The Current Study

Given the heterogeneity present across the ASD spectrum, educators continue to need access to reading intervention protocols that are feasible for use with existing school personnel and are able to address the unique and varying needs of students. Our goal was to develop an intervention to meet the specialized needs of ASD that can be delivered by teachers or paraprofessionals. This multicomponent intervention was developed over multiple years through

a federally funded iterative development process. This mixed method approach employed a combination observation data, student assessment data, and a series of single case design studies across multiple sites. The multiple iterations of the intervention were guided by our primary data collection as well as more recent empirical investigations that contributed to understanding the anomalies unique to ASD. This study was designed to test the initial efficacy or proof of concept of this set of instructional practices.

Rationale and Research Questions

The aim of this study was to investigate the effects of a multicomponent reading intervention on the reading outcomes of middle grade students with ASD. We set out to answer the following research questions:

When differences in verbal ability and fluency are controlled for, to what extent are there differences in vocabulary outcomes for students classified with different levels of ASD symptom severity between the multicomponent reading intervention condition and the comparison condition for students with ASD in grades 3-8?

When differences in verbal ability and fluency are controlled for, to what extent are there differences in reading comprehension outcomes for students classified with different levels of ASD symptom severity between the multicomponent reading intervention condition and the comparison condition for students with ASD in grades 3-8?

Method

Description of School Districts and Participants

This study took place in two separate geographic sites in near urban districts located in the south midwestern and southwestern United States. Both sites had two school districts and five schools that participated in the study. Across all districts, intervention sessions were

conducted with no other students present. Space for sessions varied and included a private conference room, empty classroom, a safe room, a section of the library, and the hallway.

South midwestern site. Nine students with ASD participated in the study from the larger of the two districts at the south-central location, which was located approximately 20 miles outside of a large metropolitan area. According to district records six students were Hispanic and three were White. The near urban district serves over 20,000 students with 14 elementary and six middle school campuses. According to district data, the student body consists of 48% Hispanic (any race), 45% White, 4% Black, and 1% Asian. One elementary school and two middle schools participated in the study. The participating elementary campus has an enrollment of 627 students and the two middle schools both had an enrollment of approximately 800 students.

Three students with ASD participated in the study from the smaller district which is considered more rural and is located approximately 30 miles from the same large metropolitan city as the larger district. According to district records two students were White and one was Asian. This district serves 4,900 students with 3 elementary schools. Seventy-three percent of the student population were white, 18% Hispanic, and 28% economically disadvantaged.

Southwestern site. Twelve students with ASD participated in the study from the larger of the two at the southwestern locations which is located approximately 25 miles outside of a large metropolitan area. Due to a change in special education administration and Covid-19 surges we were unable to acquire demographics for the participants in this district. This near urban district serves over 4,300 students in grades K-8 with four elementary schools and one middle school campus. The demographic makeup of the students attending the district includes 69.3% Hispanic, 19.6% White, 5.8% Black, 3.1% Asian, and 1% Native American. The elementary

school enrollments ranges from 573 to 900 students (Mean enrollment = 711). The middle school enrollment is 1,350 students.

Four students with ASD participated in the study from the smaller district at the southwestern location. Two students were White, one Hispanic, and one Native American. This school is a certified educational therapy charter school which provides specialized educational and behavioral services for students with low-incidence disabilities (e.g., ASD, LD) in grades K-8 serving approximately 50 students across two campus locations. The demographic makeup of the students attending the district were 31% White, 23% Hispanic, 15% Black, and 19% Asian and 12% Native American.

Description of intervention tutors. The intervention tutors were hired, trained, and supervised by senior members of the research team at both sites. At the south midwestern location, five researchers (two females and three males) employed the sponsoring research center of the university served as intervention tutors. Four researchers tutored students for all sessions. The other researcher was an as needed substitute and also managed materials and coordinated logistics. All held advanced degrees in education (1 doctorate, 3 completed master's degrees and 1 in progress); had previously held positions as classroom teachers and specialists (e.g., reading specialist, instructional coach); and had experience working with individuals with ASD.

The intervention tutors at the southwestern location were paraprofessionals from the cooperating school districts that were hired by the research team for work hours in addition to their positions with the schools. All the tutors were referred by district personnel to participant in the study due to their expertise working in the district and with students with disabilities. The six female tutors consisted of one certified teacher and 5 paraprofessionals. One intervention tutor had a master's degree. Demographic data for the remaining five intervention tutors was not

available.

Procedures

District administrators were provided a description of participants that met the following criterion: (a) school-based primary eligibility under the ASD category, (b) evidence of reading problems including not passing the grade level state reading test or an IEP goal/objective to improve reading outcomes, (c) did not participate in the state's alternative assessment, (d) not currently identified as a students with limited English proficiency. Parental consent and student assent were obtained for all participants as approved by the universities' Institutional Review Board requirements.

Students were assessed by researchers independent of the intervention team. All assessors went through extensive training including reliability checks to a gold standard. Working closely with school personnel, the pretest and posttest battery were administered based on information included on each student's individualized education plan (IEP). The assessment team was unaware of the conditions assigned to students during the pretest/posttest.

We conducted a two-group experimental study using a matched randomized design. Within each district a research methodologist independent of the assessment and intervention matched students on ASD severity (GARS) and pretest reading fluency (AimsWeb) using the Mahalanobis distance metric to map the two covariate vectors into a single number (Stuart, 2010). Distance for the Mahalanobis metric is the sum of the normalized distances for each covariate adjusted for covariance in the data. The rationale for these procedures was empirically derived from studies suggesting the influence of cognitive phenotype (Mcyntyre et al., 2017) and reading fluency (Solari et al., 2017) on reading outcomes for students with ASD. Controlling these factors within the design improved the ability to understand the impact of the intervention

on the outcomes of vocabulary and reading comprehension. Students were matched into pairs and then randomly assigned to the intervention condition (n = 15) or the comparison condition (n = 13). There reason there were 15 in the intervention and 13 in comparison is because an odd number of students were recruited in one district. Similar procedures for assignment to condition have been employed in previous reading intervention studies of students with ASD (Bailey & Ariciuli, 2017; Roux et al., 2015; Roux, Dion, & Barrette, 2015).

Tutor training. Each tutor attended a total of two half days of training prior to providing the intervention to students. The training was provided by the principal investigator at one site and the co-principal investigator at the second site. Training materials and procedures were consistent across sites. Training consisted of reviewing the following: purpose of the study, structure of the study (e.g. 1:1 tutoring sessions), each component of the intervention (e.g. vocabulary), behavior management techniques (e.g. specific feedback), features of effective instruction (e.g. explicit instruction), and audio recording sessions plus tracking procedures.

The first training session included reviewing the purpose and structure of the study, behavior management techniques, and features of effective instruction. Tutors practiced using their assigned iPod and recording. The second training session included a review of the topics presented in the first training session, reviewing each component of the intervention, intervention materials, modeling, role play, and coaching. A member from the research team modeled the procedural steps for each component and instructed the tutors to engage in role play. Tutors were placed in groups of two to practice presenting the intervention materials and procedural steps. During this time, a member from the research team walked around the room, provided coaching/specific feedback (e.g. you missed step number two, let's try it like this), modeled, and instructed tutors to continue practicing. This training session, ended with a review of all the

materials and components of the intervention, presenting a completed student binder, and answering questions. the research team provided fully developed instructional materials and lesson plans. All intervention tutors received binders of the lesson plans and student materials. In addition, Intervention tutors received a training packet that included the power point presentation, visual examples of the intervention materials, steps for audio recording sessions, and the research team's contact information. Researchers provided weekly coaching sessions for each intervention tutor to further support the initial training.

Measures

The Gilliam Autism Rating Scale, Third Edition (GARS; Gilliam, 2013). The GARS was administered during the screening procedure for each cohort of participants. The purpose of GARS administration was to provide information regarding the classification of symptom severity of ASD (Level 1, 2, 3). Level 1 indicates less symptom severity whereas Level 3 indicates students with the most symptom severity of ASD characteristics. This information was used to determine student matches prior to randomization and as part of data analysis. The GARS is a standardized assessment of social interaction and communication for individuals suspected of having ASD usually completed by the student's teacher or case manager. Internal consistency reliability coefficients for the subscales exceed 0.85 and the autism indexes exceed 0.93 (Gilliam, 2013).

AIMSweb, Oral Reading Fluency (AIMSweb, 2001). Oral reading fluency passages [Grade 4]. (Available at www.aimsweb.com). The AIMSweb ORF is a one-minute timed reading of text that is leveled by grade level. The ORF subtest was administered during pretest only for purposes of matching students into pairs prior to randomization, part of data analysis and to determine the readability levels of text used during intervention for each student.

Kaufman Brief Intelligence Test, 2nd Edition (KBIT; Kaufman & Kaufman, 2004). The KBIT verbal subtest was administered during pretest and used descriptively. The KBIT is a 15-minute individually administered measure composed of two separate scales that assess verbal and nonverbal intelligence in people from 4 through 90 years of age. The split-half reliability coefficients range from .82 to .94. The KBIT-2 verbal subtest was used in the analysis to control for verbal ability.

Woodcock-Johnson III, Passage Comprehension Subtest (WJ-PC; Woodcock, McGrew, & Mather, 2001). The WJ-PC was administered as a pretest/posttest measure of reading comprehension. The WJ-PC is a nationally normed, individually administered assessment used to assess reading comprehension with standard scores. A median split-half coefficient of .92 to .96 is reported for the WJ-PC.

The Test of Sentence Reading Efficiency and Comprehension (TOSREC; Wagner, Torgesen, Rashotte, & Pearson, 2010). The TOSREC was administered as a pre/post measure of fluency and reading comprehension. The TOSREC is a nationally normed, 3-minute, group-administered assessment used to assess reading fluency and comprehension with standard scores. The alternate form-reliability coefficients exceed .85.

Strategy Use Measure (SUM; Scammacca, 2017). The SUM was administered as a pre/post measure proximal measures of reading comprehension. The raw scores of the SUM were intended to measure students' use of two comprehension strategies: (1) question generation and (2) identifying the main idea, with data supporting reliability of scoring (Scammacca, 2017). The researcher-developed SUM part one consisted of three reading passages, each followed by the same three open-ended items. These items required the student to write one easy question about the passage; one difficult question about the passage; and the main idea, or gist, of the

passage. The research-developed SUM part two, consisted of the same three reading passages, followed by multiple choice items. This item required the student to identify the main idea of each reading passage from a set of four options. As part of a validity study, the three multiple-choice items were added based on an analysis of data from a previous study suggesting that the open-ended, main-idea item might be too difficult for some students (Scammacca, 2017). The inter-rater reliability exceeds 95%.

Vocabulary measure. The researcher-developed vocabulary measure was administered as a pretest/posttest measure to determine the treatment effect of the words directly taught. This researcher-developed measure was derived from a list of 158 words that appear on both the Academic Vocabulary List (Coxhead, 2000) and the New Academic Vocabulary List (Gardner & Davies, 2014). Two senior members of the research team selected 30 words with high academic and social skill utility. The measure consisted of several sets of words and definitions, thus requiring the students to match the vocabulary terms used within the context of a sentence. Raw scores were used in the analysis of vocabulary outcomes. The test-retest *Pearson's R* reliability was 0.75.

Alternate reading inventory (ARI; citation removed for review) The ARI captures the time of instruction, type of curriculum, frequency, duration, and group size for each student in the comparison condition. We administered the alternate reading inventory form to understand and characterize the comparison condition and its relationship to the intervention.

Social validity self-report. Following completion of the intervention, treatment participants completed a social validity self-report measure. A member of the testing team guided students through 10 questions. The questions consisted of 10 questions 9 forced-choice questions with a 4-point Likert scale and one open-ended question, which have been tested and refined

from previous studies (Citation removed for review). Participants self-report perceptions of satisfaction by expressing the degree of agreement with statements such as "I really enjoy working with a tutor during reading sessions" and "The reading sessions really help me." The open-ended questions ask about students' favorite part of the reading lessons. See Table 2 for a summary of mean scores. Student responses indicated that they favored the DIMS component followed by the vocabulary component of the intervention. Students also indicated that they enjoyed working one-on-one with the tutor and having a choice to select the reading passage during fluency instruction. Under the context of asking students with ASD to work on academic tasks that are typically very difficult to complete, we view the social validity results favorably.

Intervention

Students in the treatment group received 1:1 intervention session 4-5 times a week for 30 minutes over the course of 6-8 weeks. The intervention was administered once per day. However, the intervention could be administered up to two times a day to help make-up any missed sessions. The total number of intervention sessions ranged from 23 to 30 sessions ($M = 27$). Intervention sessions occurred at different times of the day for each participant. The specific times were pre-selected by school personnel who worked with each student's teacher to determine a time that did not interfere with critical activities and other services. All sessions were held privately in pre-identified areas including conference rooms and empty classrooms.

Intervention tutors implemented a standardized protocol approach with mechanisms to adjust readability levels of text based on the individual needs of students. The multicomponent intervention consisted of the following instructional components: (a) vocabulary instruction, (b) fluency with text, and (c) reading comprehension. These components were scaffolded with a buildup of difficulty level by first focusing on word-level understanding, then sentence-level

comprehension, and finally multi-paragraph passage reading.

Vocabulary instruction. The instructional materials included 30 vocabulary words. We adapted the form of vocabulary instruction from randomized control trial (RCT) studies on interventions for students with reading difficulties (Citations removed for review) and findings from single case design studies of reading interventions for students with ASD (Citations removed for review). On the first page of the materials the words were presented in large print along with a clear and concise definition of the word and a student-friendly visual representation of the word's meaning. The instructor materials included prompts to facilitate brief discussion to support explicit connections between the target word and the visual representation. On the next second page, the Intervention tutor and students read and discussed two to three synonyms and a sentence with context clues to support understanding of the targeted word. The intervention tutor concluded with two discussion questions or sentence stems based on individualized student need. Discussion questions were designed to elicit student's prior knowledge and application of the word related to personal experiences and to promote higher-order thinking. Sentence stems were designed as a scaffold support option for students who needed additional guidance. As such, the sentence stems were based on the content of each discussion question. We split up the presentation of materials from one to two pages based on the findings from the iterative development process. This served as a mechanism for focusing the child's attention appropriately, a challenge which is common for instructing many students with ASD.

Fluency with text. Reading fluency is a multi-dimensional construct consisting of reading rate, accuracy, and prosody (Pikulski & Chard, 2005; Wolf & Katzir-Cohen, 2001). The approach to reading fluency was partially based on the deficits with prosody in speech, which is common symptom for many individuals with ASD (Holbrook & Israelsen, 2020; McCann &

Peppe, 2003). Taking into account the unique characteristics of ASD associated with prosody and language development (Eigsti, et al., 2012) and the conceptual framework (SVR and WCC), we conceptualized an approach to fluency instruction with activities to support the multidimensional characteristics of fluency including rate, accuracy, and prosody. The first part of instruction includes a mini-lesson (3-5 min) on prosody designed to support students learning about appropriate phrasing by teaching students about grammatical cues as idea units (e.g., commas, periods, exclamation point). The intervention tutor described the grammatical cue and then provided a model of fluent reading with accurate expression and phrasing. Students listened to how the teacher read "with meaning" aligned phrasing with the punctuation. This was followed by a prompt for the student to read aloud and pay attention to the targeted punctuation. The addition of the prosody mini-lesson was based on data from the intervention development project that indicated that a large number of students with ASD do not phrase text appropriately based on grammatical cues. Following similar procedures of fluency instruction for students with ASD (e.g., Barnes & Rehfeldt, 2013; Reisener et al., 2014), the intervention tutor modeled fluent oral reading utilizing expository text from *QuickReads* passages (Hiebert, 2003) followed by guided and independent practice. We developed instructional materials for all five readability levels of the *QuickReads* passages to address the heterogeneity of the students and provide intervention tutors with materials appropriately aligned with student need. The initial word correct per minute (wcpm) scores from the AIMS Web ORF were used to determine the appropriate readability levels for each student. After repeated reading, the intervention tutor checked for understanding by asking students about the main ideas of the passage (*What is this paragraph about? What is the most important idea? Tell me the most important idea in this part? What is the main idea?).* In order to assist in the student's social skill development and engagement, they

were offered a choice to select one of two readings for each lesson.

Reading comprehension. During the iterative development process, many of the students with ASD struggled with comprehension at the multi-paragraph level. To address this, we included a sentence comprehension instruction in addition to multi-paragraph main idea summarization activities. This approach provided a scaffold based on the volume and complexity of text from the sentence level to the multi-paragraph level. The sentence-level comprehension instruction consisted of the "Does it Make Sense?" (DIMS) activity, in which students identified whether specific sentences made sense or not based on the content of the *QuickReads* fluency passages. Students used context clues to determine if the sentences made sense or not. Students circled "yes" if a statement makes sense or "no" if it did not. If the student circled no, they would underline context clues to support their answer. The DIMS activity has shown to improve reading outcomes for students with reading difficulties (Citation removed for review).

The instructional routine for the main idea summarization instruction supported a text-based approach to intervention to eliminate as many extraneous details and facilitated a focus on reading the content of the text while having meaningful discussions with the Intervention tutors. We wrote question prompts to support students identifying main ideas from these passages (*What is this paragraph about? What is the most important idea? Tell me the most important idea in this part? What is the main idea?).* When students were unable to support their answer, the intervention tutor would ask the student to reread and focus on a reduced portion of text, (i.e., single sentences, single word). From the iterative development process, we concluded that this text-based approach rather than cognitive strategy instruction facilitated more instructional time focused reading rather than explanation and discussion of strategies.

The instructional materials for the main idea summarization component included

passages adapted from Newsela (https://newsela.com), a source of readings with more challenging readability levels than the fluency passages. These passages were used to support building knowledge of current events with social studies and science content. Using Newsela passages was also advantageous because we were able to adjust the Lexile level between two different readability levels based on student need. As with the fluency with text component, we developed instructional materials for both readability levels to meet the diverse needs of students with ASD. The rationale for only two readability levels at the multi-paragraph level was the intent was to scaffold student's access to text even when it was "stretch" text with respect to their readability compared to the fluency text. This was feasible because of the extensive supports that were available in a one-on-one instructional setting. The intent of this approach was to assist students in advancing towards the types of texts they would be expected to read in the classroom and to assist them with strategies for successfully understanding more complex text structures. Several experimental and single case design studies report improvements in comprehension outcomes for students with ASD when main idea summarization instruction and questioning strategies were explicitly taught (Asberg & Sandburg, 2010; Bailey et al., 2017; Bethune & Wood, 2013; Howorth et al., 2016; Reutebuch et al., 2015; Roux, Dion, & Barrette, 2015; Roux et al., 2015; Turner, 2017; Whalon & Hanline, 2008).

Comparison Condition

Findings from the ARI indicated that students in the comparison received between 9 – 60 mins of reading instruction (M = 36 mins) for 18 to 36 weeks (M = 30 weeks) Groups ranged in size from one student to 29 students (M = 7 students, median = 3, bimodal = 1, 2). Of the students in the comparison condition, 60% received their primary reading instruction outside of general education. In terms of instructional approach, teachers reported 30% of students received

computer adaptive instruction (i.e., I-Ready), 23% received specialized instruction from the special education teacher, 17% received instruction from a phonics or fluency program (e.g., Corrective Reading Decoding, Read Naturally), with 5% receiving instruction from the general education English language arts basal. Teachers reported frequent use of independent reading and answering questions and with low level reading passages to support vocabulary development. Additional tertiary interventions were reported for 21% of students.

Fidelity of Implementation

Intervention tutors audio-recorded all the intervention sessions. The audio recordings were grouped as beginning, middle, and end of the intervention. Within each time period, one audio was randomly selected for a total of three to be coded for fidelity. A total of 30% of intervention sessions were coded. Prior to coding, a researcher participated in a 4-hour training conducted by a senior member of the research team. In adherence with the procedures for establishing a gold standard (Gwet, 2001) the senior researcher coded two audios. Gold standard procedures were followed to establish interrater reliability by having a second researcher code the same audios followed by a meeting to discuss discrepancies. After coding of the second video, the interrater agreement between the two researchers was 100%. A total adherence percentage was calculated by dividing the total number of steps that were presented by the total number of steps listed.

The implementation validity checklist (IVC) included the essential use of the instructional routines and materials of each component (i.e., Vocabulary, Fluency with text, reading comprehension). Each step of the intervention was evaluated for adherence by documenting its presence or absence. In addition, global indicators for student management and quality of implementation were assessed on a scale of one (lowest quality) to four (highest

quality). The overall global indicator mean was 3.6 (scale 1-4). The presence of instruction indicating adherence to the intervention was 91% for the coded sessions. At one site the adherence was 93% and the second site the adherence was 89%.

Data Analysis Plan

Data were analyzed using repeated measures multilevel models in HLM 7. At level-1, scores at pre-test and post-test for each outcome (i.e., vocabulary, SUM, TOSREC, and WJ-PC) were modeled as functions of student-level variables plus random error. At level-2 and 3, coefficients for student-level variables were modeled as functions of district-level averages, which in turn were modeled as functions of grand averages plus random errors for pre-test and post-test scores (i.e., residuals between district averages and grand average). To identify significant effects, we selected an alpha level of 0.10. The significance level for hypothesis testing is a value for which the *p* value less than or equal to is considered statistically significant. Typical alpha level values include 0.10, 0.05, and 0.001 (Ross, 2017). Due to the associated low statistical power to detect effects, the small sample size, and the exploratory nature of the study it was important to consider the probabilities of both type I and type II errors. With small sample size studies, the probability of type II error with a *p* value of 0.05 may be too high. In this study, setting the p value at 0.10 was appropriate under consideration of the null hypothesis being false even though the *p* value is greater than 0.05 to avoid a type II error (McCabe et al., 2017). We also calculated and reported Hedges' *g* effect sizes from the posttest descriptive statistics for the outcome measures (See Table 3). Hedges' g was selected because it provides a conservative estimate of effect when used for studies with small sample sizes (Hedges & Olkin, 1985).

Final models were determined through step-wise regression procedures. Initially, level-1 equations included coefficients/fixed effects for Baseline (i.e., pre-test scores), Time (i.e., growth

or decline in post-test scores for comparison students), Intervention exposure (i.e., growth or decline in post-test scores for intervention students, beyond that expected for comparison students), the interaction between GARS severity level and intervention exposure (i.e., dummy variables for 2 groups, which separated intervention effect estimates for the 3 GARS categories), AIMSweb scores at pre-test, K-BIT-Verbal scores at pre-test, and Age at pre-test, as well as an error term. Iteratively, models were run in HLM 7 and analysis output was inspected for insignificant coefficient estimates. One at a time, the variable with the largest p-value was removed. Models were re-run until all remaining coefficients' p-values were sufficiently low. When GARS severity level was retained in models, dummy variables were cycled to obtain estimates of effects for each severity group. See Table 3 for descriptive statistics and reported effect sizes for posttest scores of outcome measures.

Results

The multicomponent intervention was associated with significant gains in WJ-PC scores of 22.62 (*se*=8.19, *df*=3, *p*=.070) for students at GARS severity level 1. Significant gains favoring the intervention were also detected on vocabulary score of 10.19 (*se*=2.78, *df*=3, *p*=.035) and for students rated at GARS severity level 1 and for students rated at GARS severity level 2 [vocabulary score 5.46 (*se*=1.60, *df*=3, *p*=.042)]. No significant differences between the multicomponent intervention and the comparison condition were detected for SUM or TOSREC scores. We describe the results from the final models for each dependent variable.

WJ-PC Model Selection

For WJ-PC outcomes, the step-wise regression procedure involved removal of the variables: Interaction of GARS groups and Baseline, Time and Age. Explanatory variables retained in the final model included: Baseline (i.e., a single intercept), Intervention effect and

interaction with GARS severity level, AIMSweb, and K-BIT-Verbal. The variables AIMSweb and K-BIT-Verbal were centered on the grand mean, rendering other coefficient estimates representative of effects for students with average scores on the measures.

WJ-PC fixed effects. On average, students' baseline scores were 83.10 (*se*=3.72). For students rated at GARS severity level 1, intervention was associated with significant gains in WJ-PC scores of 22.62 (*se*=8.19, *df*=3, *p*=.070, δ=0.35). Changes in WJ-PC scores observed for students rated at GARS severity levels 2 and 3 were comparatively smaller and not significant (γ_{300}=8.31, *se*=4.93, *df*=3, *p*=.190; γ_{400}=-1.43, *se*=9.97, *df*=3, *p*=.895). Differences between score changes in GARS severity levels 1 and 3 were significant (γ_{200}- γ_{400}=24.05, *df*=3, *p*=.079), although other contrasts were not significant (γ_{200}- γ_{300}=14.31, *df*=3, *p*=.127; γ_{300}- γ_{400}=9.74, *df*=3, *p*=.366). See Table 4 for a summary of the results.

Vocabulary Model Selection.

For vocabulary outcomes, the step-wise regression procedure involved removal of the variables: Interaction of GARS severity levels and Baseline, Time, AIMSweb, and K-BIT-Verbal. Explanatory variables retained in the final model included: Baseline (i.e., a single intercept), Intervention effect and interaction with GARS severity level, and Age. The variable Age was centered on the grand mean, rendering other coefficient estimates representative of effects for students of the average age of 11. Because Time was excluded from the final model, the intercept coefficient Baseline captures pre-test scores for both intervention and comparison students, as well as post-test scores for comparison students.

Vocabulary fixed effects. On average, students' baseline scores were 7.21 (*se*=1.14). Respectively for students rated at GARS severity levels 1 and 2, intervention was associated with significant gains in vocabulary scores of 10.19 (*se*=2.78, *df*=3, *p*=.035, δ=1.15) and 5.46

(*se*=1.60, *df*=3, *p*=.042, δ=0.48). Gains observed for students rated at GARS severity level 3 were comparatively smaller and not significant (γ_{400}=1.69, *se*=3.40, *df*=3, *p*=.652). Differences between score increases in GARS severity level 1 and 3 were significant (γ_{200}- γ_{400}=8.50, *df*=3, *p*=.063), although other contrasts were not significant (γ_{200}- γ_{300}=4.73, *df*=3, *p*=.152; γ_{300}- γ_{400}=3.77, *df*=3, *p*=.325). See Table 5 for a summary of the results.

SUM Model Selection.

For SUM outcomes, the step-wise regression procedure involved removal of the variables: Interaction of GARS severity levels and Baseline, Time, Intervention effect, Interaction between intervention and GARS severity levels, and AIMSweb. Explanatory variables retained in the final model included: Baseline (i.e., a single intercept), K-BIT-Verbal and Age. Both variables were centered on the grand means, rendering the Baseline coefficient estimate representative of expected scores for students with average K-BIT-Verbal scores (i.e., X) and an age of 11. Because Time and Intervention effects were excluded from the final model, the intercept coefficient (i.e., Baseline) provides an estimate of both pre- and post-test scores for comparison and intervention students.

SUM fixed effects. On average (across time and groups) students' SUM scores were 8.48 (*se*=.085). SUM scores were higher for students who scored highly on the K-BIT-Verbal at pre-test. For each point above the average K-BIT-Verbal score, SUM scores were expected to be 0.20 points higher. This incremental effect was significant (*se*=0.05, *df*=27, *p*<.001). SUM scores were also higher for older students in the sample. Incremental increases of 0.96 were expected for every year of age above the average of 11 (and incremental decreases of 0.96 were expected for every year of age below 11). This effect was significant (*se*=0.49, *df*=27, *p*=.062). See Table 6 for a summary of the results.

TOSREC Model Selection.

For TOSREC outcomes, the step-wise regression procedure involved removal of the variables: Interaction of GARS severity levels and Baseline, Intervention effect, Interaction of intervention and GARS severity levels, and Age. Explanatory variables retained in the final model included: Baseline (i.e., a single intercept), Time, AIMSweb, and K-BIT-Verbal. AIMSweb and K-BIT-Verbal were centered on the grand means, rendering the coefficient estimates for Baseline and Time representative of effects expected for students with average AIMSweb scores (i.e., X) and K-BIT-Verbal scores (i.e., X). Because Intervention effect was excluded from the final model, the Time coefficient captures growth from pre- and post-test scores for both comparison and intervention students.

TOSREC fixed effects. On average students' baseline TOSREC scores were 9.13 (*se*=14.05). At post-test, students' scores increased by an average of 9.21 (*se*=2.01, *df*=3, *p*=.019). Growth in TOSREC scores was higher for students who scored highly on the AIMSweb and K-BIT-Verbal at pre-test. For each point above the average AIMSweb score, TOSREC scores were expected to be 0.22 points higher. This incremental effect was significant (*se*=0.07, *df*=27, *p*=.003). For each point above the average K-BIT-Verbal score, TOSREC scores were expected to be 0.57 points higher. This incremental effect was significant (*se*=0.17, *df*=27, *p*=.002). See Table 7 for a summary of the results.

Discussion

This experimental pilot study aimed to examine the impact of a multicomponent reading comprehension intervention for students with ASD. The results from this study show promise for improving reading outcomes for students with ASD classified with level 1 and level 2 symptom severity on the GARS. Students classified with level 1 symptom severity made statistically

significant ($p = 0.07$) gains on one standardized measure of reading comprehension and a proximal measure of vocabulary ($p = 0.035$). Students classified with level 2 symptom severity also made statistically gains ($p = 0.042$) on the proximal measure of vocabulary. These findings align with recent findings from reader profile studies which suggest that ASD symptom severity had a significant impact on reading outcomes with the most intractable reading problems existing for students with the highest symptom severity levels (Huemer & Mann, 2010; Mcyntyre et al., 2017; Solari et al., 2019). In this study, students classified as GARS level 3 symptom severity did not respond to the intervention. This may indicate that the intervention was not sufficiently intense enough in order to detect differences for students with the highest levels of symptom severity (GARS 3 level 3). This pilot study contributes to the evidence base of reading comprehension practices that are associated with impact for specific subgroups of students with ASD (Bailey & Ariciuli, 2020; Brown et al., 2013; Chiang & Lin, 2007; El Zein, et al., 2014; Finnegan & Mazin, 2016; Knight et al., 2013; Knight & Sartini, 2015; Senokossoff, 2016; Whalon et al., 2009).

Fletcher & Wagner (2014) argue that intervention research should de-emphasize the *p* values and instead should focus on small but meaningful effects that may accumulate over time. Further, the statistical significance is not related to the practical significance or the magnitude of the effect of an intervention. Other factors such as sample size, sample variance on outcome variables and the covariates included in the analysis influence statistical significance (Lipsey et al., 2012). Due to the small number of group design studies currently in the literature and the lack of studies establishing baseline reading performance over time (i.e., Hill et al., 2008) for students with ASD, it is difficult to surmise specifically how the findings from this study and others might translate to practical differences in classroom performance.

The Hedges' *g* effect sizes reported across conditions in this study for the WJ-PC scores (g = 0.20) and the vocabulary scores (*g* = 0.43) align with mean effect sizes reported from a recent meta-analysis of reading interventions for older students with reading difficulties (Scammacca et al., 2015). This also coincides with the descriptive data showing a consistent pattern of higher average performance at posttest for the intervention over the comparison condition across all the dependent measures (See Table 3).

We interpreted the findings from this pilot study as promising for several reasons. First, students with ASD represent a heterogeneous group of students whose communication, social, and academic needs vary considerably with some students demonstrating significant problems (referred to a level 3 on the GARS) and other students demonstrating less significant difficulties (referred to as level 1 on the GARS). Differentiating treatment effects for students with ASD is a valuable goal in assisting educators in determining appropriate treatments that align with the specific needs of their students. Second, considering the differential needs of students with ASD, it is of high importance to ensure that these students receive the necessary evidence-based instruction required. Third, we view this intervention as feasible for implementation based on high levels of fidelity with instruction being provided by para professionals at one site. This was possible by employing structured materials that were also capable of being adapted with differing readability levels to address the heterogeneity of performance that is typical across the ASD spectrum. The issue of feasibility is also supported in light of the ARI findings of the comparison condition indicating that many students were already receiving 1:1 intervention. With this in mind, the differences detected show promise in this approach to intervention being an improvement over current typical practice.

While significant gains were not detected on the remaining proximal and standardized dependent measures, the findings from the analysis of those outcome measures provided insights that are important for consideration. Specifically, the influence of verbal ability was detected on both measures showing that students with higher KBIT verbal scores performed better. In addition, the influence of participant age was detected for the proximal reading measure showing that older students performed better. The influence of reading fluency was detected for the second standardized measure of reading showing that students with higher baseline fluency had higher scores on this dependent measure.

By employing group-design methodology, this study provides a meaningful contribution to the literature. To date, we are currently aware of only three group-design experimental studies focused on reading interventions for students with ASD (See Table 1). Because the study utilized a matching pair with randomization allocation to condition, we were able to better understand for whom the intervention was effective. The fidelity of implementation data show that tutors in the treatment condition demonstrated close adherence to the intervention protocol.

Many students with ASD demonstrate social and communication difficulties which are manifested in their challenges in academic areas, especially with reading comprehension. Often, students with ASD are thought of as primarily having challenges that require social-behavioral support, whereas their academic needs may be under recognized (Citation removed for review). In addition to these social-emotional and communication challenges, many students with ASD demonstrate difficulties in literacy with significant reading comprehension problems that inhibit their access to content learning, reading for meaning, and knowledge. Recognizing that reading comprehension is often the gateway to future academic success in secondary and post-secondary

settings, we designed and tested an approach to improving reading comprehension for students with ASD.

Limitations

Due to the small sample size and to avoid a type II error, we set the *p* value at 0.10 rather than the more universally accepted *p* value of 0.05 which we acknowledge as a limitation to the data analysis. While the sample size is small, it is within the same magnitude of the few experimental studies including participants with ASD (i.e., Bailey et al., 2017). We recognize that because this is a pilot experimental study, a larger replication study is needed to improve confidence in the findings. We do acknowledge the reduced power associated with this sample size as well and suggest that a replication study not only attain a larger sample, but that it focuses on students with Level 1 and 2 ASD severity levels (GARS; Gilliam, 2013) to improve confidence in the findings. It is also possible that the total time of instruction was a potential confound of the study. This was unclear since data was not collected on the type or amount of reading instruction that continued to be provided by the schools for the students in the intervention condition.

Implications for Research

While we interpreted the findings as promising, future studies require both replication and refinement with larger fully powered sample sizes. From both a theoretical and empirical perspective, this study reinforces the importance of taking into account ASD symptomology and oral language to assist in isolating on differences that may be present on reading outcomes between different instructional conditions. Considering the impact of verbal ability on the non-significant dependent measures, future research should consider additional measures of language such as receptive listening skills to more broadly operationalize the construct of language. It also

may be fruitful for future studies to consider language-based interventions integrated with reading instruction. The approach of a language-based listening comprehension intervention has shown promise for younger students with ASD (Henry & Solari, 2020). Future studies should consider procedures to measure the business as usual condition and the reading instruction students in the intervention condition continue to receive in addition to the intervention

Implications for Instruction

Although growth is modest and appears to be limited to participants who have lower levels of ASD symptom severity, considering the need for evidence-based practices we interpret the findings as having implications for instruction. This multicomponent intervention distinguishes itself from more general reading support by providing instruction 1:1 and with materials that are both structured and able to be flexed with differing readability levels in order to address the heterogenous needs that are prevalent for students with ASD. The findings support the use of multicomponent instruction that includes vocabulary, fluency with text, and reading comprehension (e.g., main idea generation). The scaffold of building up the task of understanding concepts ~~from~~ in text (from word, to sentence, to multi-paragraph) and the flexibility with readability levels are two approaches for school personnel to consider. We also learned several important lessons from this pilot experimental study that might be considered in future studies.

First, the students in this study were eager to participate in rigorous academic instruction addressing vocabulary and comprehension development that was situated to their learning needs. Most students in this study indicated that they liked taking part in the intervention sessions. Students were especially complimentary of the activities involving learning new vocabulary with the visuals and graphic organizers, along with the *Does It Make* Sense and *Quick Read* activities

because they were funny and interesting. Interestingly, many students showed interest in learning about how grammatical cues can help them to understand how to approach phrasing of text. The interventionist also reported students taking the initiative to apply the prosodic lessons to the repeated reading portion of the fluency component. Second, students acknowledged liking the opportunity to work one-on-one to address their vocabulary and comprehension development with a 1:1 tutor. Of course, it is possible that this was influenced by selection bias on the part of students and parents who chose to participate in the study. Third, students recognized and articulated that the reading sessions were helpful to them. This is especially relevant considering the emphasis on 'reading to learn' in the middle grades and beyond, as well as a lack of attention focused on improving specific reading skills of older students with ASD. Participants were cognizant of their need for reading improvement and eager for the support provided, albeit, some more reluctant at first than others. We surmise from the participants' social validity data and feedback from tutors and classroom teachers, that the positive attitudes about the intervention and overall engagement of the learners had to do with the treatment's provision of a safe space for students to participate in rigorous, consistent academic instruction with a tutor skilled in addressing specific learning needs with adaptable, appropriate materials.

References

AIMSweb. (2001). Oral reading fluency passages [Grade 4]. Available at http://www.aimsweb.com.

Åsberg, J., & Sandberg A.D. (2010). Discourse comprehension intervention for high-functioning students with autism spectrum disorders: Preliminary findings from a school-based study. *Journal of Research in Special Educational Needs,10*(2), 91–98. doi:10.1111/j.1471-3802.2010.01147.x

Bailey, B., Arciuli, J. (2020). Reading instruction for children with autism spectrum disorders: A systematic review and quality analysis. *Review Journal of Autism and Developmental Disorders*,*7*, 127-150. doi: 10.1007/s40489-019-00185-8

Bailey, B., Arciuli, J., & Stancliffe, R. J. (2017). Effects of ABRACADABRA instruction on spelling in children with autism spectrum disorder. *Scientific Studies of Reading*, *21*(2), 146-164. doi:10.1080/10888438.2016.1276183

Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, *45*(3), 601-613.

Barnes, C. S., & Rehfeldt, R. A. (2013). Effects of fluency instruction on selection-based and topography-based comprehension measures. *Research in Autism Spectrum Disorders*, *7*(6), 639-647. doi: 10.1016/j.rasd.2013.02.010

Bethune, K. S., & Wood, C. L. (2013). Effects of coaching on teachers' use of function-based interventions for students with severe disabilities. *Teacher Education and Special Education*, *36*(2), 97–114. doi:10.1177/0888406413478637

Brown, H. M., Oram-Cardy, J., Johnson, A. (2013). A meta-analysis of the reading

comprehension skills of individuals on the autism spectrum. *Journal of Autism and Developmental Disorders,* 43(4), 932-955. http://dx.doi.org/10.1007/s10803-012-1638-1

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities*, *36*(2), 151–164. doi: 10.1177/002221940303600208

Centers for Disease Control and Prevention. (n.d.). Data & Statistics on Autism Spectrum Disorder. Retrieved from https://www.cdc.gov/ncbddd/autism/data.html

Chiang, H. M., & Lin, Y. H. (2007). Reading comprehension instruction for students with autism spectrum disorders: A review of the literature. *Focus on Autism and Other Developmental Disabilities*, *22*(4), 259–267. doi:10.1177/10883576070220040801

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238. doi:10.2307/3587951

Dugan, E., Kamps, D., Leonard, B., Watkins, N., Rheinberger, A., & Stackhaus, J. (1995). Effects of cooperative learning groups during social studies for students with autism and fourth□grade peers. *Journal of Applied Behavior Analysis*, *28*(2), 175-188. doi:10.1901/jaba.1995.28-175

Eigsti, I. M., Schuh, J., Mencl, E., Schultz, R. T., & Paul, R. (2012). The neural underpinnings of prosody in autism. *Child Neuropsychology*, *18*(6), 600-617 doi:10.1080/09297049.2011.639757

El Zein, F., Solis, M., Vaughn, S., & McCulley, L. (2014). Reading comprehension interventions for students with autism spectrum disorders: A synthesis of research. *Journal of Autism and Developmental Disorders*, *44*(6), 1303–1322. doi:10.1007/s10803-013-1989-2

Finnegan, E., & Mazin, A.L. (2016). Strategies for increasing reading comprehension skills in students with autism spectrum disorder: A review of the literature. *Education and Treatment of Children*, *39*(2), 187-219. doi:10.1353/etc.2016.0007

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology*, *1*(4), 329–342. doi:10.1111/j.2044-835X.1983.tb00906.x

Fletcher, J. M., & Wagner, R. K. (2014). Accumulating knowledge: When are reading intervention results meaningful? *Journal or Research on Educational Effectiveness*, 7(3), 294-299, http://dx.doi.org/10.1080/19345747.2014.925310

Ganz, J. B., & Flores, M. M. (2009). The effectiveness of direct instruction for teaching language to children with autism spectrum disorders: Identifying materials. *Journal of Autism and Developmental Disorders, 39*(1), 75–83. doi:10.1007/s10803-008-0602-6

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, *35*(3), 305–327. doi:10.1093/applin/amt015

Gilliam, J. E. (2013). *Gilliam autism rating scale*, Third Edition (GARS-3). Austin, TX: Pro-ed.

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

Grandin, T., Duffy, K., & Attwood, T. (2004). *Developing talents: Careers for people with asperger syndrome*. Shawnee Mission, KS: Autism Asperger Publishing Company.

Grindle, C. F., Carl Hughes, J., Saville, M., Huxley, K., & Hastings, R. P. (2013). Teaching early reading skills to children with autism using MimioSprout Early Reading. *Behavioral Interventions*, *28*(3), 203-224. doi:10.1002/bin.1364

Gwet, K. (2001). *Handbook of inter-rater reliability: How to estimate the level of agreement between two or multiple raters*. Gaithersburg, MD: STATAXIS Publishing Company.

Happe, F. (2005). The weak central coherence account of autism. In F. Volkmar, P. A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 640–649). Hoboken, NJ: John Wiley & Sons.

Happe, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. doi:10.1007/s10803-005-0039-0

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. San Diego, CA: Academic Press.

Henry, A. R., & Solari, E. J. (2020). Targeting oral language and listening comprehension development for students with autism spectrum disorder: A school-based pilot study. *Journal of Autism and Developmental Disorders*, *50*(10), 3763-3776. http://dx.doi.org/10.1007/s10803-020-04434-2

Hiebert, E. (2003). *QuickReads: A research-based fluency program*. Modern Curriculum Press: MA.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172-177. http://dx.doi.org/10.1111/j.1750-8606.2008.00061.x

Holbrook, S., & Israelsen, M. (2020). Speech prosody interventions for persons with autism spectrum disorders: A systematic review. *American Journal of Speech-Language Pathology, 29*(4), 2189-2205. http://dx.doi.org/10.1044/2020_AJSLP-19-00127

Howorth, S., Lopata, C., Thomeer, M., & Rodger, J. (2016). Effects of the TWA strategy on

expository reading comprehension of students with autism. *British Journal of Special Education,* 43(1), 39-59, http://dx.doi.org/10.1111/1467-8578.12122

Huemer, S. V., & Mann, V. (2010). A comprehensive profile of decoding and comprehension in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *40*(4), 485-493. doi: 10.1007/s10803-009-0892-3

Kamps, D., Locke, P., Delquadri, J., & Hall, R. V. (1989). Increasing academic skills of students with autism using fifth grade peers as tutors. *Education and Treatment of Children, 12*(1), 38-51. doi:10.1901/jaba.1994.27-49

Kamps, D. M., Barbetta, P. M., Leonard, B. R., & Delquadri, L. (1994). Classwide peer tutoring: An integration strategy to improve reading skills and promote peer interactions among students with autism and general education peers. *Journal of Applied Behavior Analysis, 27*(1), 49–61. doi:10.1901/jaba.1994.27-49

Kamps, D. M., Leonard, B., Potucek, J., & Garrison-Harrell, L. (1995). Cooperative learning groups in reading: An integration strategy for students with autism and general classroom peers. *Behavioral Disorders, 21*(1), 89–109. doi:10.1177/019874299502100103

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test*, Second Edition (K-BIT-2). Minneapolis, MN: Pearson Assessment.

Knight, V., McKissick, B. R., & Saunders, A. (2013). A review of technology-based interventions to teach academic skills to students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *43*(11), 2628-2648. doi: 10.1007/s10803-013-1848-1

Knight, V. F., & Sartini, E. (2015). A comprehensive literature review of comprehension strategies in core content areas for students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(5), 1213-1229.

doi: 10.1007/s10803-014-2280-x

Knight, V. F., Wood, C. L., Spooner, F., Browder, D. M., & O'Brien, C. P. (2015). An exploratory study using science eTexts with students with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 30*(2), 86–99. https://doi.org/10.1177/1088357614559214

Koegel, R., Koegel, L., & Carter, C. (1999). Pivotal teaching interactions for children with autism. *School Psychology Review*, *28*(4), 576–594. doi: 10.1080/02796015.1999.12085986

Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., Busick, M.D. (2012). *Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms*. (NCSER 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. This report is available on the IES website at http://ies.ed.gov/ncser/.

Lucas, R., & Norbury, C. F. (2014). Levels of text comprehension in children with autism spectrum disorders (ASD): The influence of language phenotype. *Journal of Autism and Developmental Disorders*, *44*(11), 2756-2768. doi: 10.1007/s10803-014-2133-7

Lyon, G. R. (1998). Why reading is not a natural process. *Educational Leadership, 55*(6), 14–18.

McCabe, G. P., Craig, B. A., Moore, D. S. (2017). Introduction to the Practice of Statistics. London, United Kingdom: Macmillan Learning.

McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical

review. *International Journal of Language & Communication Disorders*, *38*(4), 325-350. http://dx.doi.org/10.1080/1368282031000154204

McIntyre, N. S., Solari, E. J., Grimm, R. P., Lerro, L. E., Gonzales, J. E., & Mundy, P. C. (2017). A comprehensive examination of reading heterogeneity in students with high functioning autism: Distinct reading profiles and their relation to autism symptom severity. *Journal of Autism and Developmental Disorders*, *47*(4), 1086–1101. doi:10.1007/s10803-017-3029-0

Norbury, C., & Nation, K. (2011). Understanding variability in reading comprehension in adolescents with autism spectrum disorders: Interactions with language status and decoding skill. *Scientific Studies of Reading*, 15(3), 191–210. http://dx.doi.org/10.1080/10888431003623553

Onaolapo, A. Y., & Onaolapo, O. J. (2017). Global data on autism spectrum disorders prevalence: A review of facts, fallacies and limitations. *Universal Journal of Clinical Medicine*, *5*(2), 14-23.

Pikulski, J.J. and Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, *58*(6), 510-519. doi:10.1598/RT.58.6.2

Quill, K. A. (2000). *Do-watch-listen-say: Social and communication intervention for children with autism*. Baltimore, MD: Paul H. Brookes.

Randi, J., Newman, T., & Grigorenko, E. L. (2010). Teaching children with autism to read for meaning: Challenges and possibilities. *Journal of Autism and Developmental Disorders*, *40*(7), 890-902. doi: 10.1007/s10803-010-0938-6

Reisener, C. D., Lancaster, A. L., McMullin, W. A., & Ho, T. (2014). A preliminary investigation of evidence-based interventions to increase oral reading fluency in children with

autism. *Journal of Applied School Psychology*, *30*(1), 50-67. http://dx.doi.org/10.1080/15377903.2013.869785

Ricketts, J., Jones, C. R., Happé, F., & Charman, T. (2013). Reading comprehension in autism spectrum disorders: The role of oral language and social functioning. *Journal of Autism and Developmental Disorders*, *43*(4), 807-816. http://dx.doi.org/10.1007/s10803-012-1619-4

Ross, S. M. (2017). *Introductory Statistics*. London, United Kingdom: Academic Press.

Roux, C., Dion, E., & Barrette, A. (2015). Enhancing reading comprehension among students with high-functioning autism spectrum disorder: A randomized pilot study. *Canadian Journal of Education*, *38*(1), 1-19. doi:10.1177/0741932514533998

Roux, C., Dion, E., Barrette, A., Dupéré, V., & Fuchs, D. (2015). Efficacy of an intervention to enhance reading comprehension of students with high-functioning autism spectrum disorder. *Remedial and Special Education*, *36*(3), 131–142. doi: 10.1177/0741932514533998

Reutebuch, C. K., El Zein, F., Kim, M. K., Weinberg, A. N., & Vaughn, S. (2015). Investigating a reading comprehension intervention for high school students with autism spectrum disorder: A pilot study. *Research in Autism Spectrum Disorders*, *9*, 96–111. doi: 10.1016/j.rasd.2014.10.002

Scammacca, N. K. (2017). Technical summary: Strategy use measure. *Meadows Center for Preventing Educational Risk.*

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities*, *48*(4), 369-390. http://dx.doi.org/10.1177/0022219413504995

Senokossoff, G. W. (2016). Developing reading comprehension skills in high-functioning children with autism spectrum disorder: A review of the research, 1990–2012. *Reading & Writing Quarterly*, *32*(3), 223-246. doi: 10.1080/10573569.2014.936574

Solari, E. J., Grimm, R., McIntyre, N. S., Swain-Lerro, L., Zajic, M., & Mundy, P. C. (2017). The relation between text reading fluency and reading comprehension for students with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *41*, 8-19. doi: 10.1016/j.rasd.2017.07.002

Solari, E. J., Grimm, R. P., McIntyre, N. S., Zajic, M., & Mundy, P. C. (2019). Longitudinal stability of reading profiles in individuals with higher functioning autism. *Autism*, *23*(8), 1911–1926. doi: 10.1177/1362361318812423

Sorenson Duncan, T., Karkada, M., Deacon, S. H., & Smith, I. M. (2021). Building meaning: Meta-analysis of component skills supporting reading comprehension in children with autism spectrum disorder. *Autism Research*, 14(5), 840-858. http://dx.doi.org/10.1002/aur.2483

Stuart E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1–21. Retrieved from https://doi.org/10.1214/09-STS313

Turner, H., Remington, A., & Hill, V. (2017). Developing an intervention to improve reading comprehension for children and young people with autism spectrum disorders. *Educational & Child Psychology*, *34*(2), 13–26. Retrieved from http://discovery.ucl.ac.uk/1557365/

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). TOSREC: *Test of silent reading efficiency and comprehension*. Pro-ed.

Whalon, K. J., Al Otaiba, S., & Delano, M. E. (2009). Evidence-based reading instruction for

individuals with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *24*(1), 3–16. doi: 10.1177/1088357608328515

Whalon, K., & Hanline, M. F. (2008). Effects of a reciprocal questioning intervention on the question generation and responding of children with autism spectrum disorder. *Education and Training in Developmental Disabilities, 43*(3), 367–387. Retrieved from http://www.jstor.org/stable/23879798

Williamson, P., Carnahan, C., & Jacobs, J. (2009). *Increasing comprehension for students with autism: From research to practice*. Paper presented at the Council for Exceptional Children Conference, Seattle, WA.

Williamson, P., Carnahan, C. R., Birri, N., & Swoboda, C. (2015). Improving comprehension of narrative using character event maps for high school students with autism spectrum disorder. *The Journal of Special Education*, *49*(1), 28–38. https://doi.org/10.1177/0022466914521301

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside.

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, *5*(3), 211-239. doi: 10.1207/S1532799XSSR0503_2

Yell, M. L., & Bateman, D. F. (2017). Endrew F. v. Douglas County School District (2017) FAPE and the U.S. Supreme Court. *Teaching Exceptional Children*, *50*(1), 7–15. doi:10.1177/0040059917721116

Reading Enhancements for Students with Autism Spectrum Disorder: A Matched Randomized Pilot Experimental Study

The prevalence rates of autism spectrum disorder (ASD) across the world were recently estimated to be approximately one of every 132 persons (Baxter et al., 2015). The global rates of ASD continue to rise based on data from developed and developing countries (Onaolapo & Onaolapo, 2017). In comparison to other disabilities, the global prevalence of ASD in children is greater than conduct disorder and attention deficit hyperactive disorder (ADHD) combined (Baxter et al., 2015).

In the United States the prevalence rates of ASD for children continue to increase dramatically (1 in 88 children in 2008, 1 in 68 children in 2014, and 1 in 59 children in 2018; Center for Disease Control and Prevention, n.d.). The increased prevalence of children with ASD require the need for empirically validated interventions, specifically for the educationally high priority area of reading development. In light of the landmark Supreme Court case providing clarification regarding the interpretation of the Free and Appropriate Public Education (FAPE) clause (Yell & Bateman, 2017), school districts are now required to go above and beyond the previous held "de minimis" standard. Thus, educators need access to more specialized, evidence-based instruction as part of their specialized instructional supports required to improve academic outcomes.

For students with ASD, improving reading comprehension is critical, increasing both their opportunities for successful postsecondary experiences including attending college and obtaining meaningful employment. Higher levels of reading comprehension are associated with greater gains in other academic areas, higher levels of employment, increased independence, and overall improved quality of life (Lyon, 1998). Improved academic performance may be

just as important as social skills training for individuals with ASD (Grandin et al., 2004). This study contributes to the small yet growing body of literature investigating reading interventions designed to improve reading performance for students with ASD.

Theoretical Framework and Empirical Underpinnings

Through a multi-year iterative development process, we developed a multicomponent reading intervention derived from several theoretical and empirical lines of research: (a) theoretical underpinnings of reading instruction derived from the Simple View of Reading (SVR; Gough & Tunmer, 1986), (b) the linkage to learning challenges of many students with ASD as informed by the Weak Central Coherence cognitive processing theory (WCC; Happe & Frith, 2006; Koegel, et al., 1999; Quill, 2000), and (c) empirical findings from reader profile studies of students with ASD (Huemer & Mann, 2010; Lucas & Norbury 2014; McIntyre et al., 2017; Randi et al., 2010; Ricketts et al., 2013; Solari et al., 2017; Solari et al., 2019).

Theoretical framework. The Simple View of Reading (SVR) posits that reading comprehension is the product of word recognition and linguistic comprehension (Gough & Tunmer, 1986). Weaknesses in one or both lead to poor reading comprehension (Catts et al., 2003). There are several studies that replicate support for SVR with samples of students with ASD (Brown et al., 2013; Huemer & Mann, 2010; Lucas & Norbury 2014; McIntyre et al., 2017; Randi et al., 2010; Ricketts, 2011; Sorenson Duncan et al., 2021). For example, a meta-analysis by Brown et al., (2013) comparing neurotypical students to students with ASD reported the two strongest predictors of reading comprehension to be decoding (55% of the variance) and semantic knowledge (57% of the variance). Brown et al., (2013) concluded that taking into account language ability is essential to understanding the reading performance of students with ASD. These findings were further extended by Sorenson Duncan et al., (2021) by taking into

account the associations between several subcomponents of language (vocabulary, morphology, syntax, pragmatics, and listening comprehension) and the strength of these associations with reading comprehension. The findings from Sorensen et al., (2021) further support the importance of taking into account oral language skills as an essential consideration with reading comprehension for students with ASD.

A second theoretical model often applied to students with ASD is the cognitive processing theory of weak central coherence (WCC) (Happe & Frith, 2006; Koegel et al., 1999; Quill, 2000). WCC posits that many students with ASD have difficulty summarizing salient points, understanding main ideas, and/or overly focus on extraneous details (Happe, 2005; Williamson et al., 2009). The WCC processing theory explains ASD in terms of students' difficulties in recognizing big ideas and/or relating big ideas with details. The SVR integrated with the WCC provided an initial heuristic for designing instructional practices responding to the needs of students with ASD. The WCC cognitive processing theory integrates with both sides of the SVR in that it helps to inform decoding and linguistic development.

Empirical underpinnings – reader profile studies. While studies from 30 years ago generally supported the idea that students with ASD demonstrated profiles of high decoding and low comprehension (e.g., Frith & Snowling, 1983), more recent studies report greater heterogeneity in students' performance on word reading and comprehension measures while also accounting for additional factors such as language and other components of reading (i.e., fluency) (Solari et al., 2017). Larger sample sizes and longitudinal studies of linguistic profiles afforded opportunities for more sophisticated analyses such as the use of latent variables and model fit indices to inform interactions and make predictions (Huemer & Mann, 2010; Lucas & Norbury 2014; McIntyre et al., 2017; Randi et al., 2010; Ricketts, 2011; Solari et al., 2017; Solari

et al., 2019). The empirical findings from these studies informed the development of this multicomponent intervention with consideration of the following: (a) the heterogeneity of students with ASD in both the word reading and linguistic comprehension components of the SVR (Huemer & Mann, 2010; Mcyntyre et al., 2017; Solari et al., 2019), (b) the importance of taking into account language development (Lucas & Norbury, 2014; Mcyntyre et al., 2017; Norbury & Nation, 2011; Ricketts et al., 2013), and (c) the evidence suggesting that cognitive phenotype and social communicative factors (i.e., GARS scores), and reading fluency scores inform predictions of reading comprehension for students with ASD (Mcyntyre et al., 2017; Solari et al., 2017).

Reader profile studies of students with ASD have validated the presence of four distinct reader profiles that align with the heuristic of the SVR (Huemer & Mann, 2010; Mcyntyre et al., 2017; Solari et al., 2019), while also highlighting the diversity of performance as considerations for conceptualizing instructional needs. Further, ASD symptomology is an important factor to consider in relation to the role of language development and reading outcomes (Huemer & Mann, 2010; Mcyntyre et al., 2017; Solari et al., 2019). Ricketts et al., (2013) conducted regression analysis with a sample of older students with ASD which indicated that oral language explained unique variance in reading comprehension (Ricketts et al., 2013). Other reader profile studies have reported that language impairment of children with ASD is associated with poor performance in reading outcomes (Lucas & Norbury, 2014; Mcyntyre et al., 2017; Norbury & Nation, 2011). Therefore, it is necessary to account for the role of language within the context of reading comprehension (Lucas & Norbury, 2014; Mcyntyre et al., 2017; Norbury & Nation, 2011; Ricketts et al., 2013). It is also necessary to account for the role of ASD symptomology for other areas of reading including reading fluency (Solari et al., 2017; Solari et al., 2019). Solari et

al., (2017) reported reading fluency as an important consideration for reading comprehension for students with ASD based greater variation being observed compared to other components of reading (Solari et al., 2017). When controlling for both decoding and language, the structural equation models reported by Solari et al., (2019) suggest that reading fluency is a significant predictor of reading comprehension for students with ASD. Therefore, controlling for reading fluency affords an opportunity to isolate on the dependent variables associated with reading comprehension (i.e., vocabulary).

Reading Intervention Research of ASD

Systematic reviews of reading interventions for students with ASD conducted over the last 15 years have contributed to our understanding of the efficacy of particular instructional approaches in reading for these students (Bailey & Ariciuli, 2020; Brown et al., 2013; Chiang & Lin, 2007; El Zein, et al., 2014; Finnegan & Mazin, 2016; Knight et al., 2013; Senokossoff, 2016; Whalon et al., 2009). The majority studies represented across these systematic reviews were implemented with researchers providing the intervention. Approximately, 75% of the studies provided 1:1 instruction for the reading intervention. The literature on reading interventions for students with ASD remains underdeveloped with an over reliance on single-case design studies. However, findings from these studies do provide an initial empirical base for utilizing modeling, guided practice, and independent practice as a mechanism to support improvements in the following areas: vocabulary and graphic organizers (Dugan et al., 1995; Grindle et al., 2013; Kamps et al., 1995; Knight et al., 2015; Williamson et al., 2015), fluency with text (Barnes & Rehfeldt, 2013; Kamps et al., 1994; Kamps et al., 1989; Reisner et al., 2014), sentence comprehension, multi-paragraph comprehension, and questioning strategies (Asberg et al., 2010; Bailey et al., 2017; Bethune & Wood, 2013; Ganz & Flores, 2009;

Reutebuch et al., 2015; Roux et al., 2015; Roux, Dion, & Barrette, 2015; Turner et al., 2017; Whalon & Hanline, 2008). See Table 1 for a summary of studies of reading interventions for students with ASD in the middle grades. These studies provide empirical support for the instructional components of vocabulary, fluency and reading comprehension

It is also important to note the small number of pre-post group design studies included in these systematic reviews and the lack of student samples with older students from upper elementary and middle school grades. Three of these studies focused on younger students (average age 9 yrs.) (Bailey & Ariciuli, 2017; Roux et al., 2015; Roux, Dion, & Barrette, 2015) with only one study utilizing a sample of older students (age 13 years) (Turner et al., 2017). There are also methodological limitations to these studies. For example, Turner et al., (2017) employed a pre-post design with no comparison group which eliminates experimental control to isolate the impact of the intervention compared to maturation. Nevertheless, the remaining studies with multiple conditions did report improvements favoring the intervention conditions compared to business as usual conditions. This pilot study for students with ASD in the middle grades contributes to the literature by testing a multicomponent intervention that, while similar in some facets, is unique in terms of the makeup of the instructional components compared to previous approaches to intervention.

The Current Study

Given the heterogeneity present across the ASD spectrum, educators continue to need access to reading intervention protocols that are feasible for use with existing school personnel and are able to address the unique and varying needs of students. Our goal was to develop an intervention to meet the specialized needs of ASD that can be delivered by teachers or paraprofessionals. This multicomponent intervention was developed over multiple years through

a federally funded iterative development process. This mixed method approach employed a combination observation data, student assessment data, and a series of single case design studies across multiple sites. The multiple iterations of the intervention were guided by our primary data collection as well as more recent empirical investigations that contributed to understanding the anomalies unique to ASD. This study was designed to test the initial efficacy or proof of concept of this set of instructional practices.

Rationale and Research Questions

The aim of this study was to investigate the effects of a multicomponent reading intervention on the reading outcomes of middle grade students with ASD. We set out to answer the following research questions:

When differences in verbal ability and fluency are controlled for, to what extent are there differences in vocabulary outcomes for students classified with different levels of ASD symptom severity between the multicomponent reading intervention condition and the comparison condition for students with ASD in grades 3-8?

When differences in verbal ability and fluency are controlled for, to what extent are there differences in reading comprehension outcomes for students classified with different levels of ASD symptom severity between the multicomponent reading intervention condition and the comparison condition for students with ASD in grades 3-8?

Method

Description of School Districts and Participants

This study took place in two separate geographic sites in near urban districts located in the south midwestern and southwestern United States. Both sites had two school districts and five schools that participated in the study. Across all districts, intervention sessions were

conducted with no other students present. Space for sessions varied and included a private conference room, empty classroom, a safe room, a section of the library, and the hallway.

South midwestern site. Nine students with ASD participated in the study from the larger of the two districts at the south-central location, which was located approximately 20 miles outside of a large metropolitan area. According to district records six students were Hispanic and three were White. The near urban district serves over 20,000 students with 14 elementary and six middle school campuses. According to district data, the student body consists of 48% Hispanic (any race), 45% White, 4% Black, and 1% Asian. One elementary school and two middle schools participated in the study. The participating elementary campus has an enrollment of 627 students and the two middle schools both had an enrollment of approximately 800 students.

Three students with ASD participated in the study from the smaller district which is considered more rural and is located approximately 30 miles from the same large metropolitan city as the larger district. According to district records two students were White and one was Asian. This district serves 4,900 students with 3 elementary schools. Seventy-three percent of the student population were white, 18% Hispanic, and 28% economically disadvantaged.

Southwestern site. Twelve students with ASD participated in the study from the larger of the two at the southwestern locations which is located approximately 25 miles outside of a large metropolitan area. Due to a change in special education administration and Covid-19 surges we were unable to acquire demographics for the participants in this district. This near urban district serves over 4,300 students in grades K-8 with four elementary schools and one middle school campus. The demographic makeup of the students attending the district includes 69.3% Hispanic, 19.6% White, 5.8% Black, 3.1% Asian, and 1% Native American. The elementary

school enrollments ranges from 573 to 900 students (Mean enrollment = 711). The middle school enrollment is 1,350 students.

Four students with ASD participated in the study from the smaller district at the southwestern location. Two students were White, one Hispanic, and one Native American. This school is a certified educational therapy charter school which provides specialized educational and behavioral services for students with low-incidence disabilities (e.g., ASD, LD) in grades K-8 serving approximately 50 students across two campus locations. The demographic makeup of the students attending the district were 31% White, 23% Hispanic, 15% Black, and 19% Asian and 12% Native American.

Description of intervention tutors. The intervention tutors were hired, trained, and supervised by senior members of the research team at both sites. At the south midwestern location, five researchers (two females and three males) employed the sponsoring research center of the university served as intervention tutors. Four researchers tutored students for all sessions. The other researcher was an as needed substitute and also managed materials and coordinated logistics. All held advanced degrees in education (1 doctorate, 3 completed master's degrees and 1 in progress); had previously held positions as classroom teachers and specialists (e.g., reading specialist, instructional coach); and had experience working with individuals with ASD.

The intervention tutors at the southwestern location were paraprofessionals from the cooperating school districts that were hired by the research team for work hours in addition to their positions with the schools. All the tutors were referred by district personnel to participant in the study due to their expertise working in the district and with students with disabilities. The six female tutors consisted of one certified teacher and 5 paraprofessionals. One intervention tutor had a master's degree. Demographic data for the remaining five intervention tutors was not

available.

Procedures

District administrators were provided a description of participants that met the following criterion: (a) school-based primary eligibility under the ASD category, (b) evidence of reading problems including not passing the grade level state reading test or an IEP goal/objective to improve reading outcomes, (c) did not participate in the state's alternative assessment, (d) not currently identified as a students with limited English proficiency. Parental consent and student assent were obtained for all participants as approved by the universities' Institutional Review Board requirements.

Students were assessed by researchers independent of the intervention team. All assessors went through extensive training including reliability checks to a gold standard. Working closely with school personnel, the pretest and posttest battery were administered based on information included on each student's individualized education plan (IEP). The assessment team was unaware of the conditions assigned to students during the pretest/posttest.

We conducted a two-group experimental study using a matched randomized design. Within each district a research methodologist independent of the assessment and intervention matched students on ASD severity (GARS) and pretest reading fluency (AimsWeb) using the Mahalanobis distance metric to map the two covariate vectors into a single number (Stuart, 2010). Distance for the Mahalanobis metric is the sum of the normalized distances for each covariate adjusted for covariance in the data. The rationale for these procedures was empirically derived from studies suggesting the influence of cognitive phenotype (Mcyntyre et al., 2017) and reading fluency (Solari et al., 2017) on reading outcomes for students with ASD. Controlling these factors within the design improved the ability to understand the impact of the intervention

on the outcomes of vocabulary and reading comprehension. Students were matched into pairs and then randomly assigned to the intervention condition (n = 15) or the comparison condition (n = 13). There reason there were 15 in the intervention and 13 in comparison is because an odd number of students were recruited in one district. Similar procedures for assignment to condition have been employed in previous reading intervention studies of students with ASD (Bailey & Ariciuli, 2017; Roux et al., 2015; Roux, Dion, & Barrette, 2015).

Tutor training. Each tutor attended a total of two half days of training prior to providing the intervention to students. The training was provided by the principal investigator at one site and the co-principal investigator at the second site. Training materials and procedures were consistent across sites. Training consisted of reviewing the following: purpose of the study, structure of the study (e.g. 1:1 tutoring sessions), each component of the intervention (e.g. vocabulary), behavior management techniques (e.g. specific feedback), features of effective instruction (e.g. explicit instruction), and audio recording sessions plus tracking procedures.

The first training session included reviewing the purpose and structure of the study, behavior management techniques, and features of effective instruction. Tutors practiced using their assigned iPod and recording. The second training session included a review of the topics presented in the first training session, reviewing each component of the intervention, intervention materials, modeling, role play, and coaching. A member from the research team modeled the procedural steps for each component and instructed the tutors to engage in role play. Tutors were placed in groups of two to practice presenting the intervention materials and procedural steps. During this time, a member from the research team walked around the room, provided coaching/specific feedback (e.g. you missed step number two, let's try it like this), modeled, and instructed tutors to continue practicing. This training session, ended with a review of all the

materials and components of the intervention, presenting a completed student binder, and answering questions. the research team provided fully developed instructional materials and lesson plans. All intervention tutors received binders of the lesson plans and student materials. In addition, Intervention tutors received a training packet that included the power point presentation, visual examples of the intervention materials, steps for audio recording sessions, and the research team's contact information. Researchers provided weekly coaching sessions for each intervention tutor to further support the initial training.

Measures

The Gilliam Autism Rating Scale, Third Edition (GARS; Gilliam, 2013). The GARS was administered during the screening procedure for each cohort of participants. The purpose of GARS administration was to provide information regarding the classification of symptom severity of ASD (Level 1, 2, 3). Level 1 indicates less symptom severity whereas Level 3 indicates students with the most symptom severity of ASD characteristics. This information was used to determine student matches prior to randomization and as part of data analysis. The GARS is a standardized assessment of social interaction and communication for individuals suspected of having ASD usually completed by the student's teacher or case manager. Internal consistency reliability coefficients for the subscales exceed 0.85 and the autism indexes exceed 0.93 (Gilliam, 2013).

AIMSweb, Oral Reading Fluency (AIMSweb, 2001). Oral reading fluency passages [Grade 4]. (Available at www.aimsweb.com). The AIMSweb ORF is a one-minute timed reading of text that is leveled by grade level. The ORF subtest was administered during pretest only for purposes of matching students into pairs prior to randomization, part of data analysis and to determine the readability levels of text used during intervention for each student.

Kaufman Brief Intelligence Test, 2nd Edition (KBIT; Kaufman & Kaufman, 2004). The KBIT verbal subtest was administered during pretest and used descriptively. The KBIT is a 15-minute individually administered measure composed of two separate scales that assess verbal and nonverbal intelligence in people from 4 through 90 years of age. The split-half reliability coefficients range from .82 to .94. The KBIT-2 verbal subtest was used in the analysis to control for verbal ability.

Woodcock-Johnson III, Passage Comprehension Subtest (WJ-PC; Woodcock, McGrew, & Mather, 2001). The WJ-PC was administered as a pretest/posttest measure of reading comprehension. The WJ-PC is a nationally normed, individually administered assessment used to assess reading comprehension with standard scores. A median split-half coefficient of .92 to .96 is reported for the WJ-PC.

The Test of Sentence Reading Efficiency and Comprehension (TOSREC; Wagner, Torgesen, Rashotte, & Pearson, 2010). The TOSREC was administered as a pre/post measure of fluency and reading comprehension. The TOSREC is a nationally normed, 3-minute, group-administered assessment used to assess reading fluency and comprehension with standard scores. The alternate form-reliability coefficients exceed .85.

Strategy Use Measure (SUM; Scammacca, 2017). The SUM was administered as a pre/post measure proximal measures of reading comprehension. The raw scores of the SUM were intended to measure students' use of two comprehension strategies: (1) question generation and (2) identifying the main idea, with data supporting reliability of scoring (Scammacca, 2017). The researcher-developed SUM part one consisted of three reading passages, each followed by the same three open-ended items. These items required the student to write one easy question about the passage; one difficult question about the passage; and the main idea, or gist, of the

passage. The research-developed SUM part two, consisted of the same three reading passages, followed by multiple choice items. This item required the student to identify the main idea of each reading passage from a set of four options. As part of a validity study, the three multiple-choice items were added based on an analysis of data from a previous study suggesting that the open-ended, main-idea item might be too difficult for some students (Scammacca, 2017). The inter-rater reliability exceeds 95%.

Vocabulary measure. The researcher-developed vocabulary measure was administered as a pretest/posttest measure to determine the treatment effect of the words directly taught. This researcher-developed measure was derived from a list of 158 words that appear on both the Academic Vocabulary List (Coxhead, 2000) and the New Academic Vocabulary List (Gardner & Davies, 2014). Two senior members of the research team selected 30 words with high academic and social skill utility. The measure consisted of several sets of words and definitions, thus requiring the students to match the vocabulary terms used within the context of a sentence. Raw scores were used in the analysis of vocabulary outcomes. The test-retest *Pearson's R* reliability was 0.75.

Alternate reading inventory (ARI; citation removed for review) The ARI captures the time of instruction, type of curriculum, frequency, duration, and group size for each student in the comparison condition. We administered the alternate reading inventory form to understand and characterize the comparison condition and its relationship to the intervention.

Social validity self-report. Following completion of the intervention, treatment participants completed a social validity self-report measure. A member of the testing team guided students through 10 questions. The questions consisted of 10 questions 9 forced-choice questions with a 4-point Likert scale and one open-ended question, which have been tested and refined

from previous studies (Citation removed for review). Participants self-report perceptions of satisfaction by expressing the degree of agreement with statements such as "I really enjoy working with a tutor during reading sessions" and "The reading sessions really help me." The open-ended questions ask about students' favorite part of the reading lessons. See Table 2 for a summary of mean scores. Student responses indicated that they favored the DIMS component followed by the vocabulary component of the intervention. Students also indicated that they enjoyed working one-on-one with the tutor and having a choice to select the reading passage during fluency instruction. Under the context of asking students with ASD to work on academic tasks that are typically very difficult to complete, we view the social validity results favorably.

Intervention

Students in the treatment group received 1:1 intervention session 4-5 times a week for 30 minutes over the course of 6-8 weeks. The intervention was administered once per day. However, the intervention could be administered up to two times a day to help make-up any missed sessions. The total number of intervention sessions ranged from 23 to 30 sessions ($M = 27$). Intervention sessions occurred at different times of the day for each participant. The specific times were pre-selected by school personnel who worked with each student's teacher to determine a time that did not interfere with critical activities and other services. All sessions were held privately in pre-identified areas including conference rooms and empty classrooms.

Intervention tutors implemented a standardized protocol approach with mechanisms to adjust readability levels of text based on the individual needs of students. The multicomponent intervention consisted of the following instructional components: (a) vocabulary instruction, (b) fluency with text, and (c) reading comprehension. These components were scaffolded with a buildup of difficulty level by first focusing on word-level understanding, then sentence-level

comprehension, and finally multi-paragraph passage reading.

Vocabulary instruction. The instructional materials included 30 vocabulary words. We adapted the form of vocabulary instruction from randomized control trial (RCT) studies on interventions for students with reading difficulties (Citations removed for review) and findings from single case design studies of reading interventions for students with ASD (Citations removed for review). On the first page of the materials the words were presented in large print along with a clear and concise definition of the word and a student-friendly visual representation of the word's meaning. The instructor materials included prompts to facilitate brief discussion to support explicit connections between the target word and the visual representation. On the next second page, the Intervention tutor and students read and discussed two to three synonyms and a sentence with context clues to support understanding of the targeted word. The intervention tutor concluded with two discussion questions or sentence stems based on individualized student need. Discussion questions were designed to elicit student's prior knowledge and application of the word related to personal experiences and to promote higher-order thinking. Sentence stems were designed as a scaffold support option for students who needed additional guidance. As such, the sentence stems were based on the content of each discussion question. We split up the presentation of materials from one to two pages based on the findings from the iterative development process. This served as a mechanism for focusing the child's attention appropriately, a challenge which is common for instructing many students with ASD.

Fluency with text. Reading fluency is a multi-dimensional construct consisting of reading rate, accuracy, and prosody (Pikulski & Chard, 2005; Wolf & Katzir-Cohen, 2001). The approach to reading fluency was partially based on the deficits with prosody in speech, which is common symptom for many individuals with ASD (Holbrook & Israelsen, 2020; McCann &

Peppe, 2003). Taking into account the unique characteristics of ASD associated with prosody and language development (Eigsti, et al., 2012) and the conceptual framework (SVR and WCC), we conceptualized an approach to fluency instruction with activities to support the multidimensional characteristics of fluency including rate, accuracy, and prosody. The first part of instruction includes a mini-lesson (3-5 min) on prosody designed to support students learning about appropriate phrasing by teaching students about grammatical cues as idea units (e.g., commas, periods, exclamation point). The intervention tutor described the grammatical cue and then provided a model of fluent reading with accurate expression and phrasing. Students listened to how the teacher read "with meaning" aligned phrasing with the punctuation. This was followed by a prompt for the student to read aloud and pay attention to the targeted punctuation. The addition of the prosody mini-lesson was based on data from the intervention development project that indicated that a large number of students with ASD do not phrase text appropriately based on grammatical cues. Following similar procedures of fluency instruction for students with ASD (e.g., Barnes & Rehfeldt, 2013; Reisener et al., 2014), the intervention tutor modeled fluent oral reading utilizing expository text from *QuickReads* passages (Hiebert, 2003) followed by guided and independent practice. We developed instructional materials for all five readability levels of the *QuickReads* passages to address the heterogeneity of the students and provide intervention tutors with materials appropriately aligned with student need. The initial word correct per minute (wcpm) scores from the AIMS Web ORF were used to determine the appropriate readability levels for each student. After repeated reading, the intervention tutor checked for understanding by asking students about the main ideas of the passage (*What is this paragraph about? What is the most important idea? Tell me the most important idea in this part? What is the main idea?).* In order to assist in the student's social skill development and engagement, they

were offered a choice to select one of two readings for each lesson.

Reading comprehension. During the iterative development process, many of the students with ASD struggled with comprehension at the multi-paragraph level. To address this, we included a sentence comprehension instruction in addition to multi-paragraph main idea summarization activities. This approach provided a scaffold based on the volume and complexity of text from the sentence level to the multi-paragraph level. The sentence-level comprehension instruction consisted of the "Does it Make Sense?" (DIMS) activity, in which students identified whether specific sentences made sense or not based on the content of the *QuickReads* fluency passages. Students used context clues to determine if the sentences made sense or not. Students circled "yes" if a statement makes sense or "no" if it did not. If the student circled no, they would underline context clues to support their answer. The DIMS activity has shown to improve reading outcomes for students with reading difficulties (Citation removed for review).

The instructional routine for the main idea summarization instruction supported a text-based approach to intervention to eliminate as many extraneous details and facilitated a focus on reading the content of the text while having meaningful discussions with the Intervention tutors. We wrote question prompts to support students identifying main ideas from these passages (*What is this paragraph about? What is the most important idea? Tell me the most important idea in this part? What is the main idea?).* When students were unable to support their answer, the intervention tutor would ask the student to reread and focus on a reduced portion of text, (i.e., single sentences, single word). From the iterative development process, we concluded that this text-based approach rather than cognitive strategy instruction facilitated more instructional time focused reading rather than explanation and discussion of strategies.

The instructional materials for the main idea summarization component included

passages adapted from Newsela (https://newsela.com), a source of readings with more challenging readability levels than the fluency passages. These passages were used to support building knowledge of current events with social studies and science content. Using Newsela passages was also advantageous because we were able to adjust the Lexile level between two different readability levels based on student need. As with the fluency with text component, we developed instructional materials for both readability levels to meet the diverse needs of students with ASD. The rationale for only two readability levels at the multi-paragraph level was the intent was to scaffold student's access to text even when it was "stretch" text with respect to their readability compared to the fluency text. This was feasible because of the extensive supports that were available in a one-on-one instructional setting. The intent of this approach was to assist students in advancing towards the types of texts they would be expected to read in the classroom and to assist them with strategies for successfully understanding more complex text structures. Several experimental and single case design studies report improvements in comprehension outcomes for students with ASD when main idea summarization instruction and questioning strategies were explicitly taught (Asberg & Sandburg, 2010; Bailey et al., 2017; Bethune & Wood, 2013; Howorth et al., 2016; Reutebuch et al., 2015; Roux, Dion, & Barrette, 2015; Roux et al., 2015; Turner, 2017; Whalon & Hanline, 2008).

Comparison Condition

Findings from the ARI indicated that students in the comparison received between 9 – 60 mins of reading instruction (M = 36 mins) for 18 to 36 weeks (M = 30 weeks) Groups ranged in size from one student to 29 students (M = 7 students, median = 3, bimodal = 1, 2). Of the students in the comparison condition, 60% received their primary reading instruction outside of general education. In terms of instructional approach, teachers reported 30% of students received

computer adaptive instruction (i.e., I-Ready), 23% received specialized instruction from the special education teacher, 17% received instruction from a phonics or fluency program (e.g., Corrective Reading Decoding, Read Naturally), with 5% receiving instruction from the general education English language arts basal. Teachers reported frequent use of independent reading and answering questions and with low level reading passages to support vocabulary development. Additional tertiary interventions were reported for 21% of students.

Fidelity of Implementation

Intervention tutors audio-recorded all the intervention sessions. The audio recordings were grouped as beginning, middle, and end of the intervention. Within each time period, one audio was randomly selected for a total of three to be coded for fidelity. A total of 30% of intervention sessions were coded. Prior to coding, a researcher participated in a 4-hour training conducted by a senior member of the research team. In adherence with the procedures for establishing a gold standard (Gwet, 2001) the senior researcher coded two audios. Gold standard procedures were followed to establish interrater reliability by having a second researcher code the same audios followed by a meeting to discuss discrepancies. After coding of the second video, the interrater agreement between the two researchers was 100%. A total adherence percentage was calculated by dividing the total number of steps that were presented by the total number of steps listed.

The implementation validity checklist (IVC) included the essential use of the instructional routines and materials of each component (i.e., Vocabulary, Fluency with text, reading comprehension). Each step of the intervention was evaluated for adherence by documenting its presence or absence. In addition, global indicators for student management and quality of implementation were assessed on a scale of one (lowest quality) to four (highest

quality). The overall global indicator mean was 3.6 (scale 1-4). The presence of instruction indicating adherence to the intervention was 91% for the coded sessions. At one site the adherence was 93% and the second site the adherence was 89%.

Data Analysis Plan

Data were analyzed using repeated measures multilevel models in HLM 7. At level-1, scores at pre-test and post-test for each outcome (i.e., vocabulary, SUM, TOSREC, and WJ-PC) were modeled as functions of student-level variables plus random error. At level-2 and 3, coefficients for student-level variables were modeled as functions of district-level averages, which in turn were modeled as functions of grand averages plus random errors for pre-test and post-test scores (i.e., residuals between district averages and grand average). To identify significant effects, we selected an alpha level of 0.10. The significance level for hypothesis testing is a value for which the *p* value less than or equal to is considered statistically significant. Typical alpha level values include 0.10, 0.05, and 0.001 (Ross, 2017). Due to the associated low statistical power to detect effects, the small sample size, and the exploratory nature of the study it was important to consider the probabilities of both type I and type II errors. With small sample size studies, the probability of type II error with a *p* value of 0.05 may be too high. In this study, setting the p value at 0.10 was appropriate under consideration of the null hypothesis being false even though the *p* value is greater than 0.05 to avoid a type II error (McCabe et al., 2017). We also calculated and reported Hedges' *g* effect sizes from the posttest descriptive statistics for the outcome measures (See Table 3). Hedges' g was selected because it provides a conservative estimate of effect when used for studies with small sample sizes (Hedges & Olkin, 1985).

Final models were determined through step-wise regression procedures. Initially, level-1 equations included coefficients/fixed effects for Baseline (i.e., pre-test scores), Time (i.e., growth

or decline in post-test scores for comparison students), Intervention exposure (i.e., growth or decline in post-test scores for intervention students, beyond that expected for comparison students), the interaction between GARS severity level and intervention exposure (i.e., dummy variables for 2 groups, which separated intervention effect estimates for the 3 GARS categories), AIMSweb scores at pre-test, K-BIT-Verbal scores at pre-test, and Age at pre-test, as well as an error term. Iteratively, models were run in HLM 7 and analysis output was inspected for insignificant coefficient estimates. One at a time, the variable with the largest p-value was removed. Models were re-run until all remaining coefficients' p-values were sufficiently low. When GARS severity level was retained in models, dummy variables were cycled to obtain estimates of effects for each severity group. See Table 3 for descriptive statistics and reported effect sizes for posttest scores of outcome measures.

Results

The multicomponent intervention was associated with significant gains in WJ-PC scores of 22.62 (*se*=8.19, *df*=3, *p*=.070) for students at GARS severity level 1. Significant gains favoring the intervention were also detected on vocabulary score of 10.19 (*se*=2.78, *df*=3, *p*=.035) and for students rated at GARS severity level 1 and for students rated at GARS severity level 2 [vocabulary score 5.46 (*se*=1.60, *df*=3, *p*=.042)]. No significant differences between the multicomponent intervention and the comparison condition were detected for SUM or TOSREC scores. We describe the results from the final models for each dependent variable.

WJ-PC Model Selection

For WJ-PC outcomes, the step-wise regression procedure involved removal of the variables: Interaction of GARS groups and Baseline, Time and Age. Explanatory variables retained in the final model included: Baseline (i.e., a single intercept), Intervention effect and

interaction with GARS severity level, AIMSweb, and K-BIT-Verbal. The variables AIMSweb and K-BIT-Verbal were centered on the grand mean, rendering other coefficient estimates representative of effects for students with average scores on the measures.

WJ-PC fixed effects. On average, students' baseline scores were 83.10 (*se*=3.72). For students rated at GARS severity level 1, intervention was associated with significant gains in WJ-PC scores of 22.62 (*se*=8.19, *df*=3, *p*=.070, δ=0.35). Changes in WJ-PC scores observed for students rated at GARS severity levels 2 and 3 were comparatively smaller and not significant (γ_{300}=8.31, *se*=4.93, *df*=3, *p*=.190; γ_{400}=-1.43, *se*=9.97, *df*=3, *p*=.895). Differences between score changes in GARS severity levels 1 and 3 were significant (γ_{200}- γ_{400}=24.05, *df*=3, *p*=.079), although other contrasts were not significant (γ_{200}- γ_{300}=14.31, *df*=3, *p*=.127; γ_{300}- γ_{400}=9.74, *df*=3, *p*=.366). See Table 4 for a summary of the results.

Vocabulary Model Selection.

For vocabulary outcomes, the step-wise regression procedure involved removal of the variables: Interaction of GARS severity levels and Baseline, Time, AIMSweb, and K-BIT-Verbal. Explanatory variables retained in the final model included: Baseline (i.e., a single intercept), Intervention effect and interaction with GARS severity level, and Age. The variable Age was centered on the grand mean, rendering other coefficient estimates representative of effects for students of the average age of 11. Because Time was excluded from the final model, the intercept coefficient Baseline captures pre-test scores for both intervention and comparison students, as well as post-test scores for comparison students.

Vocabulary fixed effects. On average, students' baseline scores were 7.21 (*se*=1.14). Respectively for students rated at GARS severity levels 1 and 2, intervention was associated with significant gains in vocabulary scores of 10.19 (*se*=2.78, *df*=3, *p*=.035, δ=1.15) and 5.46

(*se*=1.60, *df*=3, *p*=.042, δ=0.48). Gains observed for students rated at GARS severity level 3 were comparatively smaller and not significant (γ_{400}=1.69, *se*=3.40, *df*=3, *p*=.652). Differences between score increases in GARS severity level 1 and 3 were significant (γ_{200}- γ_{400}=8.50, *df*=3, *p*=.063), although other contrasts were not significant (γ_{200}- γ_{300}=4.73, *df*=3, *p*=.152; γ_{300}- γ_{400}=3.77, *df*=3, *p*=.325). See Table 5 for a summary of the results.

SUM Model Selection.

For SUM outcomes, the step-wise regression procedure involved removal of the variables: Interaction of GARS severity levels and Baseline, Time, Intervention effect, Interaction between intervention and GARS severity levels, and AIMSweb. Explanatory variables retained in the final model included: Baseline (i.e., a single intercept), K-BIT-Verbal and Age. Both variables were centered on the grand means, rendering the Baseline coefficient estimate representative of expected scores for students with average K-BIT-Verbal scores (i.e., X) and an age of 11. Because Time and Intervention effects were excluded from the final model, the intercept coefficient (i.e., Baseline) provides an estimate of both pre- and post-test scores for comparison and intervention students.

SUM fixed effects. On average (across time and groups) students' SUM scores were 8.48 (*se*=.085). SUM scores were higher for students who scored highly on the K-BIT-Verbal at pre-test. For each point above the average K-BIT-Verbal score, SUM scores were expected to be 0.20 points higher. This incremental effect was significant (*se*=0.05, *df*=27, *p*<.001). SUM scores were also higher for older students in the sample. Incremental increases of 0.96 were expected for every year of age above the average of 11 (and incremental decreases of 0.96 were expected for every year of age below 11). This effect was significant (*se*=0.49, *df*=27, *p*=.062). See Table 6 for a summary of the results.

TOSREC Model Selection.

For TOSREC outcomes, the step-wise regression procedure involved removal of the variables: Interaction of GARS severity levels and Baseline, Intervention effect, Interaction of intervention and GARS severity levels, and Age. Explanatory variables retained in the final model included: Baseline (i.e., a single intercept), Time, AIMSweb, and K-BIT-Verbal. AIMSweb and K-BIT-Verbal were centered on the grand means, rendering the coefficient estimates for Baseline and Time representative of effects expected for students with average AIMSweb scores (i.e., X) and K-BIT-Verbal scores (i.e., X). Because Intervention effect was excluded from the final model, the Time coefficient captures growth from pre- and post-test scores for both comparison and intervention students.

TOSREC fixed effects. On average students' baseline TOSREC scores were 9.13 (*se*=14.05). At post-test, students' scores increased by an average of 9.21 (*se*=2.01, *df*=3, *p*=.019). Growth in TOSREC scores was higher for students who scored highly on the AIMSweb and K-BIT-Verbal at pre-test. For each point above the average AIMSweb score, TOSREC scores were expected to be 0.22 points higher. This incremental effect was significant (*se*=0.07, *df*=27, *p*=.003). For each point above the average K-BIT-Verbal score, TOSREC scores were expected to be 0.57 points higher. This incremental effect was significant (*se*=0.17, *df*=27, *p*=.002). See Table 7 for a summary of the results.

Discussion

This experimental pilot study aimed to examine the impact of a multicomponent reading comprehension intervention for students with ASD. The results from this study show promise for improving reading outcomes for students with ASD classified with level 1 and level 2 symptom severity on the GARS. Students classified with level 1 symptom severity made statistically

significant ($p = 0.07$) gains on one standardized measure of reading comprehension and a proximal measure of vocabulary ($p = 0.035$). Students classified with level 2 symptom severity also made statistically gains ($p = 0.042$) on the proximal measure of vocabulary. These findings align with recent findings from reader profile studies which suggest that ASD symptom severity had a significant impact on reading outcomes with the most intractable reading problems existing for students with the highest symptom severity levels (Huemer & Mann, 2010; Mcyntyre et al., 2017; Solari et al., 2019). In this study, students classified as GARS level 3 symptom severity did not respond to the intervention. This may indicate that the intervention was not sufficiently intense enough in order to detect differences for students with the highest levels of symptom severity (GARS 3 level 3). This pilot study contributes to the evidence base of reading comprehension practices that are associated with impact for specific subgroups of students with ASD (Bailey & Ariciuli, 2020; Brown et al., 2013; Chiang & Lin, 2007; El Zein, et al., 2014; Finnegan & Mazin, 2016; Knight et al., 2013; Knight & Sartini, 2015; Senokossoff, 2016; Whalon et al., 2009).

Fletcher & Wagner (2014) argue that intervention research should de-emphasize the *p* values and instead should focus on small but meaningful effects that may accumulate over time. Further, the statistical significance is not related to the practical significance or the magnitude of the effect of an intervention. Other factors such as sample size, sample variance on outcome variables and the covariates included in the analysis influence statistical significance (Lipsey et al., 2012). Due to the small number of group design studies currently in the literature and the lack of studies establishing baseline reading performance over time (i.e., Hill et al., 2008) for students with ASD, it is difficult to surmise specifically how the findings from this study and others might translate to practical differences in classroom performance.

The Hedges' *g* effect sizes reported across conditions in this study for the WJ-PC scores (g = 0.20) and the vocabulary scores (*g* = 0.43) align with mean effect sizes reported from a recent meta-analysis of reading interventions for older students with reading difficulties (Scammacca et al., 2015). This also coincides with the descriptive data showing a consistent pattern of higher average performance at posttest for the intervention over the comparison condition across all the dependent measures (See Table 3).

We interpreted the findings from this pilot study as promising for several reasons. First, students with ASD represent a heterogeneous group of students whose communication, social, and academic needs vary considerably with some students demonstrating significant problems (referred to a level 3 on the GARS) and other students demonstrating less significant difficulties (referred to as level 1 on the GARS). Differentiating treatment effects for students with ASD is a valuable goal in assisting educators in determining appropriate treatments that align with the specific needs of their students. Second, considering the differential needs of students with ASD, it is of high importance to ensure that these students receive the necessary evidence-based instruction required. Third, we view this intervention as feasible for implementation based on high levels of fidelity with instruction being provided by para professionals at one site. This was possible by employing structured materials that were also capable of being adapted with differing readability levels to address the heterogeneity of performance that is typical across the ASD spectrum. The issue of feasibility is also supported in light of the ARI findings of the comparison condition indicating that many students were already receiving 1:1 intervention. With this in mind, the differences detected show promise in this approach to intervention being an improvement over current typical practice.

While significant gains were not detected on the remaining proximal and standardized dependent measures, the findings from the analysis of those outcome measures provided insights that are important for consideration. Specifically, the influence of verbal ability was detected on both measures showing that students with higher KBIT verbal scores performed better. In addition, the influence of participant age was detected for the proximal reading measure showing that older students performed better. The influence of reading fluency was detected for the second standardized measure of reading showing that students with higher baseline fluency had higher scores on this dependent measure.

By employing group-design methodology, this study provides a meaningful contribution to the literature. To date, we are currently aware of only three group-design experimental studies focused on reading interventions for students with ASD (See Table 1). Because the study utilized a matching pair with randomization allocation to condition, we were able to better understand for whom the intervention was effective. The fidelity of implementation data show that tutors in the treatment condition demonstrated close adherence to the intervention protocol.

Many students with ASD demonstrate social and communication difficulties which are manifested in their challenges in academic areas, especially with reading comprehension. Often, students with ASD are thought of as primarily having challenges that require social-behavioral support, whereas their academic needs may be under recognized (Citation removed for review). In addition to these social-emotional and communication challenges, many students with ASD demonstrate difficulties in literacy with significant reading comprehension problems that inhibit their access to content learning, reading for meaning, and knowledge. Recognizing that reading comprehension is often the gateway to future academic success in secondary and post-secondary

settings, we designed and tested an approach to improving reading comprehension for students with ASD.

Limitations

Due to the small sample size and to avoid a type II error, we set the *p* value at 0.10 rather than the more universally accepted *p* value of 0.05 which we acknowledge as a limitation to the data analysis. While the sample size is small, it is within the same magnitude of the few experimental studies including participants with ASD (i.e., Bailey et al., 2017). We recognize that because this is a pilot experimental study, a larger replication study is needed to improve confidence in the findings. We do acknowledge the reduced power associated with this sample size as well and suggest that a replication study not only attain a larger sample, but that it focuses on students with Level 1 and 2 ASD severity levels (GARS; Gilliam, 2013) to improve confidence in the findings. It is also possible that the total time of instruction was a potential confound of the study. This was unclear since data was not collected on the type or amount of reading instruction that continued to be provided by the schools for the students in the intervention condition.

Implications for Research

While we interpreted the findings as promising, future studies require both replication and refinement with larger fully powered sample sizes. From both a theoretical and empirical perspective, this study reinforces the importance of taking into account ASD symptomology and oral language to assist in isolating on differences that may be present on reading outcomes between different instructional conditions. Considering the impact of verbal ability on the non-significant dependent measures, future research should consider additional measures of language such as receptive listening skills to more broadly operationalize the construct of language. It also

may be fruitful for future studies to consider language-based interventions integrated with reading instruction. The approach of a language-based listening comprehension intervention has shown promise for younger students with ASD (Henry & Solari, 2020). Future studies should consider procedures to measure the business as usual condition and the reading instruction students in the intervention condition continue to receive in addition to the intervention

Implications for Instruction

Although growth is modest and appears to be limited to participants who have lower levels of ASD symptom severity, considering the need for evidence-based practices we interpret the findings as having implications for instruction. This multicomponent intervention distinguishes itself from more general reading support by providing instruction 1:1 and with materials that are both structured and able to be flexed with differing readability levels in order to address the heterogenous needs that are prevalent for students with ASD. The findings support the use of multicomponent instruction that includes vocabulary, fluency with text, and reading comprehension (e.g., main idea generation). The scaffold of building up the task of understanding concepts ~~from~~ in text (from word, to sentence, to multi-paragraph) and the flexibility with readability levels are two approaches for school personnel to consider. We also learned several important lessons from this pilot experimental study that might be considered in future studies.

First, the students in this study were eager to participate in rigorous academic instruction addressing vocabulary and comprehension development that was situated to their learning needs. Most students in this study indicated that they liked taking part in the intervention sessions. Students were especially complimentary of the activities involving learning new vocabulary with the visuals and graphic organizers, along with the *Does It Make* Sense and *Quick Read* activities

because they were funny and interesting. Interestingly, many students showed interest in learning about how grammatical cues can help them to understand how to approach phrasing of text. The interventionist also reported students taking the initiative to apply the prosodic lessons to the repeated reading portion of the fluency component. Second, students acknowledged liking the opportunity to work one-on-one to address their vocabulary and comprehension development with a 1:1 tutor. Of course, it is possible that this was influenced by selection bias on the part of students and parents who chose to participate in the study. Third, students recognized and articulated that the reading sessions were helpful to them. This is especially relevant considering the emphasis on 'reading to learn' in the middle grades and beyond, as well as a lack of attention focused on improving specific reading skills of older students with ASD. Participants were cognizant of their need for reading improvement and eager for the support provided, albeit, some more reluctant at first than others. We surmise from the participants' social validity data and feedback from tutors and classroom teachers, that the positive attitudes about the intervention and overall engagement of the learners had to do with the treatment's provision of a safe space for students to participate in rigorous, consistent academic instruction with a tutor skilled in addressing specific learning needs with adaptable, appropriate materials.

References

AIMSweb. (2001). Oral reading fluency passages [Grade 4]. Available at http://www.aimsweb.com.

Åsberg, J., & Sandberg A.D. (2010). Discourse comprehension intervention for high-functioning students with autism spectrum disorders: Preliminary findings from a school-based study. *Journal of Research in Special Educational Needs,10*(2), 91–98. doi:10.1111/j.1471-3802.2010.01147.x

Bailey, B., Arciuli, J. (2020). Reading instruction for children with autism spectrum disorders: A systematic review and quality analysis. *Review Journal of Autism and Developmental Disorders*,*7*, 127-150. doi: 10.1007/s40489-019-00185-8

Bailey, B., Arciuli, J., & Stancliffe, R. J. (2017). Effects of ABRACADABRA instruction on spelling in children with autism spectrum disorder. *Scientific Studies of Reading*, *21*(2), 146-164. doi:10.1080/10888438.2016.1276183

Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, *45*(3), 601-613.

Barnes, C. S., & Rehfeldt, R. A. (2013). Effects of fluency instruction on selection-based and topography-based comprehension measures. *Research in Autism Spectrum Disorders*, *7*(6), 639-647. doi: 10.1016/j.rasd.2013.02.010

Bethune, K. S., & Wood, C. L. (2013). Effects of coaching on teachers' use of function-based interventions for students with severe disabilities. *Teacher Education and Special Education*, *36*(2), 97–114. doi:10.1177/0888406413478637

Brown, H. M., Oram-Cardy, J., Johnson, A. (2013). A meta-analysis of the reading

comprehension skills of individuals on the autism spectrum. *Journal of Autism and Developmental Disorders,* 43(4), 932-955. http://dx.doi.org/10.1007/s10803-012-1638-1

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities*, *36*(2), 151–164. doi: 10.1177/002221940303600208

Centers for Disease Control and Prevention. (n.d.). Data & Statistics on Autism Spectrum Disorder. Retrieved from https://www.cdc.gov/ncbddd/autism/data.html

Chiang, H. M., & Lin, Y. H. (2007). Reading comprehension instruction for students with autism spectrum disorders: A review of the literature. *Focus on Autism and Other Developmental Disabilities*, *22*(4), 259–267. doi:10.1177/10883576070220040801

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238. doi:10.2307/3587951

Dugan, E., Kamps, D., Leonard, B., Watkins, N., Rheinberger, A., & Stackhaus, J. (1995). Effects of cooperative learning groups during social studies for students with autism and fourth□grade peers. *Journal of Applied Behavior Analysis*, *28*(2), 175-188. doi:10.1901/jaba.1995.28-175

Eigsti, I. M., Schuh, J., Mencl, E., Schultz, R. T., & Paul, R. (2012). The neural underpinnings of prosody in autism. *Child Neuropsychology*, *18*(6), 600-617 doi:10.1080/09297049.2011.639757

El Zein, F., Solis, M., Vaughn, S., & McCulley, L. (2014). Reading comprehension interventions for students with autism spectrum disorders: A synthesis of research. *Journal of Autism and Developmental Disorders*, *44*(6), 1303–1322. doi:10.1007/s10803-013-1989-2

Finnegan, E., & Mazin, A.L. (2016). Strategies for increasing reading comprehension skills in students with autism spectrum disorder: A review of the literature. *Education and Treatment of Children*, *39*(2), 187-219. doi:10.1353/etc.2016.0007

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology*, *1*(4), 329–342. doi:10.1111/j.2044-835X.1983.tb00906.x

Fletcher, J. M., & Wagner, R. K. (2014). Accumulating knowledge: When are reading intervention results meaningful? *Journal or Research on Educational Effectiveness*, 7(3), 294-299, http://dx.doi.org/10.1080/19345747.2014.925310

Ganz, J. B., & Flores, M. M. (2009). The effectiveness of direct instruction for teaching language to children with autism spectrum disorders: Identifying materials. *Journal of Autism and Developmental Disorders, 39*(1), 75–83. doi:10.1007/s10803-008-0602-6

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, *35*(3), 305–327. doi:10.1093/applin/amt015

Gilliam, J. E. (2013). *Gilliam autism rating scale*, Third Edition (GARS-3). Austin, TX: Pro-ed.

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

Grandin, T., Duffy, K., & Attwood, T. (2004). *Developing talents: Careers for people with asperger syndrome*. Shawnee Mission, KS: Autism Asperger Publishing Company.

Grindle, C. F., Carl Hughes, J., Saville, M., Huxley, K., & Hastings, R. P. (2013). Teaching early reading skills to children with autism using MimioSprout Early Reading. *Behavioral Interventions*, *28*(3), 203-224. doi:10.1002/bin.1364

Gwet, K. (2001). *Handbook of inter-rater reliability: How to estimate the level of agreement between two or multiple raters*. Gaithersburg, MD: STATAXIS Publishing Company.

Happe, F. (2005). The weak central coherence account of autism. In F. Volkmar, P. A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 640–649). Hoboken, NJ: John Wiley & Sons.

Happe, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. doi:10.1007/s10803-005-0039-0

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. San Diego, CA: Academic Press.

Henry, A. R., & Solari, E. J. (2020). Targeting oral language and listening comprehension development for students with autism spectrum disorder: A school-based pilot study. *Journal of Autism and Developmental Disorders*, *50*(10), 3763-3776. http://dx.doi.org/10.1007/s10803-020-04434-2

Hiebert, E. (2003). *QuickReads: A research-based fluency program*. Modern Curriculum Press: MA.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172-177. http://dx.doi.org/10.1111/j.1750-8606.2008.00061.x

Holbrook, S., & Israelsen, M. (2020). Speech prosody interventions for persons with autism spectrum disorders: A systematic review. *American Journal of Speech-Language Pathology, 29*(4), 2189-2205. http://dx.doi.org/10.1044/2020_AJSLP-19-00127

Howorth, S., Lopata, C., Thomeer, M., & Rodger, J. (2016). Effects of the TWA strategy on

expository reading comprehension of students with autism. *British Journal of Special Education,* 43(1), 39-59, http://dx.doi.org/10.1111/1467-8578.12122

Huemer, S. V., & Mann, V. (2010). A comprehensive profile of decoding and comprehension in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *40*(4), 485-493. doi: 10.1007/s10803-009-0892-3

Kamps, D., Locke, P., Delquadri, J., & Hall, R. V. (1989). Increasing academic skills of students with autism using fifth grade peers as tutors. *Education and Treatment of Children, 12*(1), 38-51. doi:10.1901/jaba.1994.27-49

Kamps, D. M., Barbetta, P. M., Leonard, B. R., & Delquadri, L. (1994). Classwide peer tutoring: An integration strategy to improve reading skills and promote peer interactions among students with autism and general education peers. *Journal of Applied Behavior Analysis, 27*(1), 49–61. doi:10.1901/jaba.1994.27-49

Kamps, D. M., Leonard, B., Potucek, J., & Garrison-Harrell, L. (1995). Cooperative learning groups in reading: An integration strategy for students with autism and general classroom peers. *Behavioral Disorders, 21*(1), 89–109. doi:10.1177/019874299502100103

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test*, Second Edition (K-BIT-2). Minneapolis, MN: Pearson Assessment.

Knight, V., McKissick, B. R., & Saunders, A. (2013). A review of technology-based interventions to teach academic skills to students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *43*(11), 2628-2648. doi: 10.1007/s10803-013-1848-1

Knight, V. F., & Sartini, E. (2015). A comprehensive literature review of comprehension strategies in core content areas for students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(5), 1213-1229.

doi: 10.1007/s10803-014-2280-x

Knight, V. F., Wood, C. L., Spooner, F., Browder, D. M., & O'Brien, C. P. (2015). An exploratory study using science eTexts with students with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 30*(2), 86–99. https://doi.org/10.1177/1088357614559214

Koegel, R., Koegel, L., & Carter, C. (1999). Pivotal teaching interactions for children with autism. *School Psychology Review*, *28*(4), 576–594. doi: 10.1080/02796015.1999.12085986

Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., Busick, M.D. (2012). *Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms*. (NCSER 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. This report is available on the IES website at http://ies.ed.gov/ncser/.

Lucas, R., & Norbury, C. F. (2014). Levels of text comprehension in children with autism spectrum disorders (ASD): The influence of language phenotype. *Journal of Autism and Developmental Disorders*, *44*(11), 2756-2768. doi: 10.1007/s10803-014-2133-7

Lyon, G. R. (1998). Why reading is not a natural process. *Educational Leadership, 55*(6), 14–18.

McCabe, G. P., Craig, B. A., Moore, D. S. (2017). Introduction to the Practice of Statistics. London, United Kingdom: Macmillan Learning.

McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical

review. *International Journal of Language & Communication Disorders*, *38*(4), 325-350. http://dx.doi.org/10.1080/1368282031000154204

McIntyre, N. S., Solari, E. J., Grimm, R. P., Lerro, L. E., Gonzales, J. E., & Mundy, P. C. (2017). A comprehensive examination of reading heterogeneity in students with high functioning autism: Distinct reading profiles and their relation to autism symptom severity. *Journal of Autism and Developmental Disorders*, *47*(4), 1086–1101. doi:10.1007/s10803-017-3029-0

Norbury, C., & Nation, K. (2011). Understanding variability in reading comprehension in adolescents with autism spectrum disorders: Interactions with language status and decoding skill. *Scientific Studies of Reading*, 15(3), 191–210. http://dx.doi.org/10.1080/10888431003623553

Onaolapo, A. Y., & Onaolapo, O. J. (2017). Global data on autism spectrum disorders prevalence: A review of facts, fallacies and limitations. *Universal Journal of Clinical Medicine*, *5*(2), 14-23.

Pikulski, J.J. and Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, *58*(6), 510-519. doi:10.1598/RT.58.6.2

Quill, K. A. (2000). *Do-watch-listen-say: Social and communication intervention for children with autism*. Baltimore, MD: Paul H. Brookes.

Randi, J., Newman, T., & Grigorenko, E. L. (2010). Teaching children with autism to read for meaning: Challenges and possibilities. *Journal of Autism and Developmental Disorders*, *40*(7), 890-902. doi: 10.1007/s10803-010-0938-6

Reisener, C. D., Lancaster, A. L., McMullin, W. A., & Ho, T. (2014). A preliminary investigation of evidence-based interventions to increase oral reading fluency in children with

autism. *Journal of Applied School Psychology*, *30*(1), 50-67. http://dx.doi.org/10.1080/15377903.2013.869785

Ricketts, J., Jones, C. R., Happé, F., & Charman, T. (2013). Reading comprehension in autism spectrum disorders: The role of oral language and social functioning. *Journal of Autism and Developmental Disorders*, *43*(4), 807-816. http://dx.doi.org/10.1007/s10803-012-1619-4

Ross, S. M. (2017). *Introductory Statistics*. London, United Kingdom: Academic Press.

Roux, C., Dion, E., & Barrette, A. (2015). Enhancing reading comprehension among students with high-functioning autism spectrum disorder: A randomized pilot study. *Canadian Journal of Education*, *38*(1), 1-19. doi:10.1177/0741932514533998

Roux, C., Dion, E., Barrette, A., Dupéré, V., & Fuchs, D. (2015). Efficacy of an intervention to enhance reading comprehension of students with high-functioning autism spectrum disorder. *Remedial and Special Education*, *36*(3), 131–142. doi: 10.1177/0741932514533998

Reutebuch, C. K., El Zein, F., Kim, M. K., Weinberg, A. N., & Vaughn, S. (2015). Investigating a reading comprehension intervention for high school students with autism spectrum disorder: A pilot study. *Research in Autism Spectrum Disorders*, *9*, 96–111. doi: 10.1016/j.rasd.2014.10.002

Scammacca, N. K. (2017). Technical summary: Strategy use measure. *Meadows Center for Preventing Educational Risk.*

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities*, *48*(4), 369-390. http://dx.doi.org/10.1177/0022219413504995

Senokossoff, G. W. (2016). Developing reading comprehension skills in high-functioning children with autism spectrum disorder: A review of the research, 1990–2012. *Reading & Writing Quarterly*, *32*(3), 223-246. doi: 10.1080/10573569.2014.936574

Solari, E. J., Grimm, R., McIntyre, N. S., Swain-Lerro, L., Zajic, M., & Mundy, P. C. (2017). The relation between text reading fluency and reading comprehension for students with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *41*, 8-19. doi: 10.1016/j.rasd.2017.07.002

Solari, E. J., Grimm, R. P., McIntyre, N. S., Zajic, M., & Mundy, P. C. (2019). Longitudinal stability of reading profiles in individuals with higher functioning autism. *Autism*, *23*(8), 1911–1926. doi: 10.1177/1362361318812423

Sorenson Duncan, T., Karkada, M., Deacon, S. H., & Smith, I. M. (2021). Building meaning: Meta-analysis of component skills supporting reading comprehension in children with autism spectrum disorder. *Autism Research*, 14(5), 840-858. http://dx.doi.org/10.1002/aur.2483

Stuart E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1–21. Retrieved from https://doi.org/10.1214/09-STS313

Turner, H., Remington, A., & Hill, V. (2017). Developing an intervention to improve reading comprehension for children and young people with autism spectrum disorders. *Educational & Child Psychology*, *34*(2), 13–26. Retrieved from http://discovery.ucl.ac.uk/1557365/

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). TOSREC: *Test of silent reading efficiency and comprehension*. Pro-ed.

Whalon, K. J., Al Otaiba, S., & Delano, M. E. (2009). Evidence-based reading instruction for

individuals with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *24*(1), 3–16. doi: 10.1177/1088357608328515

Whalon, K., & Hanline, M. F. (2008). Effects of a reciprocal questioning intervention on the question generation and responding of children with autism spectrum disorder. *Education and Training in Developmental Disabilities, 43*(3), 367–387. Retrieved from http://www.jstor.org/stable/23879798

Williamson, P., Carnahan, C., & Jacobs, J. (2009). *Increasing comprehension for students with autism: From research to practice*. Paper presented at the Council for Exceptional Children Conference, Seattle, WA.

Williamson, P., Carnahan, C. R., Birri, N., & Swoboda, C. (2015). Improving comprehension of narrative using character event maps for high school students with autism spectrum disorder. *The Journal of Special Education*, *49*(1), 28–38. https://doi.org/10.1177/0022466914521301

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside.

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, *5*(3), 211-239. doi: 10.1207/S1532799XSSR0503_2

Yell, M. L., & Bateman, D. F. (2017). Endrew F. v. Douglas County School District (2017) FAPE and the U.S. Supreme Court. *Teaching Exceptional Children*, *50*(1), 7–15. doi:10.1177/0040059917721116

Disclosure Statement

All the authors of this article declare that they have no conflict of interest.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Review Board approvals of the sponsoring institutions.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

Credit Author Statement

Author Note

Michael Solis, Graduate School of Education, University of California Riverside; Colleen K. Reutebuch, Meadows Center for Preventing Educational Risk, The University of Texas Austin; Sharon Vaughn, Department of Special Education, The University of Texas Austin; Zaira Jimenez, University of California Riverside

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A160299 to the University of California Riverside. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

Correspondence concerning this article should be addressed to Michael Solis, 1207 Sproul Hall, University of California Riverside, Riverside, CA 92521. E-mail: msolis@ucr.edu

Table 1

Empirical Underpinnings of reading interventions for students with ASD

Instructional Target (Authors)	Study Design	N	Measures	Outcomes	Implementer (group size)
Vocabulary					
Dugan et al., 1995	SCD – ABAB design	2	Pre-posttest of items learned	Mean knowledge scores increased 39%	Paraprofesional (1:1)
Grindle et al., 2013	SCD – pre-post design	3	Pre-posttest of WRAPS	Mean knowledge scores improved 47.5%	ABA tutor (1:1)
Kamps et al., 1995	SCD- reversal design	3	Pre-posttest - vocabulary	Mean academic gains scores improved 23.6%	Peer tutor (1:1)
Knight et al., 2014	SCD – multiple probe	4	Vocabulary CBMs	Mean vocabulary scores improved by 43.4%	Researcher (1:1)
Williamson et al., 2015	SCD – multiple baseline	3	Inferential knowledge	Mean knowledge scores increased 39.3%	Sped teacher (1:3)
Fluency with text					
Barnes & Rehfeldt, 2013	SCD – multiple probe	3	Percent correct- sentence-level questions	Mean percent correct scores increased 20.2%	Researcher (1:1)
Kamps et al., 1994	SCD – multiple baseline	3	Comprehension questions, wcpm	Mean scores on questions increased 28.5%, mean scores on wcpm increased by 19.7%	Peer tutor (1:1)
Kamps et al., 1989	SCD – multiple baseline	2	wcpm	Mean scores on wcpm increased 48.1%.	Peer tutor (1:1)
Reisner et al., 2014	SCD – withdrawal	4	wcpm	Mean words correct per minute for repeated reading increased 51%	Sped teacher (1:1)
Reading comprehension					
Asberg et al., 2010	Single group	12	DCT	Improvement in DCT scores, Pre-post, *ES* = 0.35	SPED teacher (small groups)
Bailey et al., 2017	Pre-post control group design	20	NARA-3 -Reading Comprehension	Improvement in passage level comprehension, T vs. C, *ES* = 0.32,	Researcher (1:1)

Instructional Target (Authors)	Study Design	N	Measures	Outcomes	Implementer (group size)
Bethune & Wood, 2013	SCD – multiple baseline	3	Literal and inferential questions	Mean scores on questions increased by 96%	Researcher (1:1)
Ganz & Flores, 2009	SCD – multiple baseline	2	Sentence-level inferential questions	Mean scores on inferential questions increased by 73.5%	Researcher (1:1)
Reutebuch et al., 2015	SCD – multiple baseline	3	Comprehension questions	Mean scores on comprehension questions increased by 14%	Peer tutor (1:1)
Turner et al., 2017	Pre-post control group design	31	YARC	Improvement in comprehension scores, T vs. C, *ES* = 1.57	Researcher (1:3)
Whalon & Hanline, 2008	SCD – multiple baseline	3	Question generation and response	Mean question generation PND = 78%, Mean response PND = 80.3%	Researcher (1:1)
Multi-component					
Roux et al., 2015	Pre-post control group design	45	Vocabulary Definitions, Main idea identification	Improvement in vocabulary scores, T vs C, *ES* = 1.0, Improvement in main idea identification scores, T vs. C, *ES* = .89	Researcher (1:3)
Roux, Dion, & Barrette, 2015	Pre-post control group design	13	Vocabulary Definitions, Main idea identification	Improvement in vocabulary scores, T vs C, *ES* = 1.3, Improvement in main idea identification scores, T vs. C, *ES* = .54	Researcher (1:3)

Note. DCT = Discourse Comprehension Test, CBMs = Curriculum-based measurement, *ES* = effect size (calculated as Hedges' g), N = number of participants; PND = percent nonoverlapping data, SCD = single case design, WJ-PC = Woodcock-Johnson Passage Comprehension Test, WRAPS = Word Recognition and Phonics Skills Test, YARC = York Assessment of Reading Comprehension, wcpm = words correct per min.,

[a.]Percent increase scores = grand mean final treatment score minus grand mean baseline score divided by points possible

Table 2

Social Validity Mean Scores

Social Validity Question	Mean scores (scale 1-5)
I really enjoyed the reading sessions.	2.80
The reading sessions really helped me.	3.33
I really enjoyed learning new vocabulary words.	3.00
I really enjoyed focusing on different kinds of punctuation.	2.60
I really enjoyed getting to choose which Quick Read to read every day.	3.33
I really enjoyed working on prosody with the Quick reads.	2.87
I really enjoyed the Does it Make Sense activities.	3.53
I really enjoyed the Main Idea activities.	2.73
I really enjoyed working with the tutor alone.	3.33

Table 3

Participant Characteristics and Pre-post Descriptive Statistics

	Group								
Variable	**Control**				**Intervention**				
N	13				15				
District 1	2				2				
District 2	6				6				
District 3	4				5				
District 4	1				2				
GARS severity level 1	2				4				
GARS severity level 2	8				9				
GARS severity level 3	3				2				
	Mean	**(SD)**			**Mean**	**(SD)**			
Age	11.08	(1.44)			11.13	(1.60)			
GARS severity scale	84.9	(17.15)			83.5	(15.32)			
AIMSWEB ORF	107.31	(36.84)			112.53	(49.43)			
K-BIT-VA	85.15	(14.81)			84.20	(17.79)			
	Pre-test		**Post-test**		**Pre-test**		**Post-test**		***ES***
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	
WJ-PC	80.69	(18.08)	86.92	(16.09)	73.73	(22.15)	90.80	(20.88)	0.20
Vocabulary	7.38	(6.16)	8.77	(7.32)	7.07	(6.92)	12.6	(9.51)	0.43
SUM	7.77	(5.95)	7.31	(5.88)	9.73	(5.18)	10.40	(5.57)	0.53
TOSREC	80.77	(19.26)	90.38	(20.98)	84.27	(19.16)	92.80	(23.71)	0.10

Note. GARS = Gilliam Autism Rating Scale, AIMS WEB ORF = Arizona Instrument to Measure Standards Oral Reading Fluency, KBIT-VA = Kaufmann Brief Intelligence Test Verbal Ability subtest, WJ-PC = Woodcock Johnson Passage Comprehension subtest, SUM = Strategy Use Measure, TOSREC = Test of Sentence Reading Efficiency and Comprehension, *ES* = effect size, T vs C.

Table 4

Woodcock Johnson Passage Comprehension Results

Dependent measure	Fixed effects (independent variables)		Model term	Coefficient estimate	Standard error	*df*	p-value	p-values for contrasts between effects for GARS groups	
WJ-PC		Baseline	γ_{000}	83.10	3.72				
		Time	-	-	-	-	*ns*		
	Intervention effects by GARS severity level	GARS level 1	γ_{200}	22.62	8.19	3	.070	1 vs. 2	.127
		GARS level 2	γ_{300}	8.31	4.93	3	.190	2 vs. 3	.366
		GARS level 3	γ_{400}	-1.43	9.97	3	.895	3 vs. 1	.079
	Effect of competencies at pre-test	AIMSWEB	γ_{500}	0.13	0.06	27	.034		
		K-BIT-Verbal	γ_{600}	0.52	0.13	27	<.001		
		Age	-	-	-	-	*ns*		

Random effects (variance terms)		Model term	Standard deviation	Variance component	df	p-value
District level	Baseline	u_{00}	8.36	69.95	3	<.001
	Time	-	-	-	-	-
	Intervention effect	u_{20}	7.74	59.98	3	<.001
Student level	Residual	e	13.64	185.98		

Note. GARS = Gilliam Autism Rating Scale, AIMS WEB ORF = Arizona Instrument to Measure Standards Oral Reading Fluency, KBIT-VA = Kaufmann Brief Intelligence Test Verbal Ability subtest, WJ-PC = Woodcock Johnson Passage Comprehension subtest

Table 5

Proximal Vocabulary Measure Results

Dependent measure	Fixed effects (independent variables)		Model term	Coefficient estimate	Standard error	*df*	p-value	p-values for contrasts between effects for GARS groups	
Vocabulary		Baseline	γ_{000}	7.21	1.14				
		Time	-	-	-	-	*ns*		
	Intervention effects by GARS severity level	GARS level 1	γ_{200}	10.19	2.78	3	.035	1 vs. 2	.152
		GARS level 2	γ_{300}	5.46	1.60	3	.042	2 vs. 3	.325
		GARS level 3	γ_{400}	1.69	3.40	3	.652	3 vs. 1	.063
	Effect of competencies at pre-test	AIMSWEB	-	-	-	-	*ns*		
		K-BIT-Verbal	-	-	-	-	*ns*		
		Age	γ_{700}	1.31	0.77	27	.102		
	Random effects (variance terms)		Model term	Standard deviation	Variance component	df	p-value		
	District level	Baseline	u_{00}	1.94	3.78	3	.021		
		Time	-	-	-				
		Intervention effect	u_{20}	3.97	15.73	3	<.001		
	Student level	Residual	e	6.65	44.26				

Note. GARS = Gilliam Autism Rating Scale, AIMS WEB ORF = Arizona Instrument to Measure Standards Oral Reading Fluency, KBIT-VA = Kaufmann Brief Intelligence Test Verbal Ability subtest.

Table 6

Strategy Use Measure Comprehension Results

Dependent measure	Fixed effects (independent variables)		Model term	Coefficient estimate	Standard error	*df*	p-value	p-values for contrasts between effects for GARS groups	
SUM		Baseline	γ_{000}	8.48	0.85				
		Time	-	-	-	-	*ns*		
								1 vs. 2	*ns*
		Intervention effect	-	-	-	-	*ns*	2 vs. 3 3 vs. 1	*ns* *ns*
	Effect of competencies at pre-test	AIMSWEB	-	-	-	-	*ns*		
		K-BIT-Verbal	γ_{600}	0.20	0.05	27	<.001		
		Age	γ_{700}	0.96	0.49	27	.062		
	Random effects (variance terms)		Model term	Standard deviation	Variance component	df	p-value		
	District level	Baseline	u_{00}	1.62	2.63	3	<.001		
		Time	-	-	-	-	-		
		Intervention effect	-	-	-	-	-		
	Student level	Residual	e	3.97	15.73				

Note. GARS = Gilliam Autism Rating Scale, AIMS WEB ORF = Arizona Instrument to Measure Standards Oral Reading Fluency, KBIT-VA = Kaufmann Brief Intelligence Test Verbal Ability subtest, SUM = Strategy Use Measure

Table 7

Test of Sentence Reading Efficiency and Comprehension Results

Dependent measure	Fixed effects (independent variables)		Model term	Coefficient estimate	Standard error	*df*	p-value	p-values for contrasts between effects for GARS groups	
TOSREC		Baseline	γ_{000}	9.13	14.05				
		Time	γ_{100}	9.21	2.01	3	.019		
								1 vs. 2	*ns*
		Intervention effect	γ_{200}	-	-	-	*ns*	2 vs. 3 3 vs. 1	*ns* *ns*
	Effect of competencies at pre-test	AIMSWEB	γ_{300}	0.22	0.07	27	.003		
		K-BIT-Verbal	γ_{400}	0.57	0.17	27	.002		
		Age	-	-	-	-	*ns*		
	Random effects (variance terms)		Model term	Standard deviation	Variance component	df	p-value		
	District level	Baseline	u_{00}	6.77	45.88	3	<.001		
		Time	u_{10}	3.41	11.65	3	<.001		
		Intervention effect	-	-	-	-	-		
	Student level	Residual	e	13.21	174.74				

Note. GARS = Gilliam Autism Rating Scale, AIMS WEB ORF = Arizona Instrument to Measure Standards Oral Reading Fluency, KBIT-VA = Kaufmann Brief Intelligence Test Verbal Ability subtest, TOSREC = Test of Sentence Reading Efficiency and Comprehension