This is the accepted version of an article published as McGregor, K. K., Van Horne, A. O., Curran, M., Cook, S. W., & Cole, R. (2021). The challenge of rich vocabulary instruction for children with developmental language disorder. *Language Speech and Hearing Services in Schools*. It is available online at https://doi.org/10.1044/2020_LSHSS-20-00110 and is made available here with the permission of the American Speech Language Hearing Association.

The Challenge of Rich Vocabulary Instruction for Children with Developmental Language Disorder

Accepted version to be published in Language, Speech, and Hearing Services in Schools

Karla K. McGregor
Center for Childhood Deafness, Language and Learning, Boys Town National Research Hospital and the Department of Communication Sciences and Disorders, University of Iowa

Amanda Owen Van Horne
Department of Communication Sciences and Disorders, University of Delaware

Maura Curran
Department of Communication Sciences and Disorders, MGH Institute of Health Professions

Susan Wagner Cook
Department of Psychological and Brain Sciences, University of Iowa

Renee Cole
Department of Chemistry, University of Iowa

Contact Author
Karla K. McGregor
Lied Learning and Technology Center
555 N. 30th St.
Omaha, NE 68131
Karla.mcgregor@boystown.org
319-471-2335
Conflict of Interest: none
Funding: NSF grant # 1748298, Karla McGregor, P.I.; Amanda Owen Van Horne, Susan Wagner Cook, Renee Cole, Co-PIs.

Abstract

Purpose: To explore responses of children with developmental language disorders (DLD) to rich vocabulary instruction and to identify potential factors that contribute to outcomes.

Method: Children with DLD participated in a language intervention embedded within a science camp. Using parent and clinician reports, standardized tests, probes, notes, and video, we derived descriptions of seven of the campers who received a vocabulary intervention that incorporated principles of rich instruction. We present them here as a case series.

Results: Five cases responded to the intervention with modest gains in Tier 2 science vocabulary and science knowledge. One case demonstrated no response and another was unable to complete the intervention. The latter two cases presented with triple risks: DLD, executive function deficits, and stressors associated with poverty. In comparison, the best responder also lived in poverty and had DLD but he had intact executive function, strengths in extant vocabulary, stronger knowledge of science, better engagement in the science and language intervention activities, and he was older. Other factors that seemed to contribute to outcomes included the complexity of the word forms and dosage.

Conclusions: Translating research on rich instruction to clinical practice is challenging. This case series motivated hypotheses about the nature of the challenge and what to do about it, the primary one being that the modest success of rich vocabulary instruction for children with DLD is not a limitation of the approach itself but, rather a reflection of the difficulty of delivering the intervention while tailoring the targets, approach, and dosage to the needs of individual children with DLD.

Key words: developmental language disorder, specific language impairment, vocabulary instruction, language intervention

The Challenge of Rich Vocabulary Instruction for Children with Developmental Language Disorder

"Closely related to text complexity and inextricably connected to reading comprehension is a focus on academic vocabulary: words that appear in a variety of content areas (such as *ignite* and *commit*). The standards call for students to grow their vocabularies through a mix of conversation, direct instruction, and reading. They ask students to determine word meanings, appreciate the nuances of words, and steadily expand their range of words and phrases."

--Common Core State Standards Initiative (2020)

It is widely recognized that age-appropriate levels of vocabulary knowledge are foundational for skilled reading (Ehri et al., 2001) and writing (Dockrell et al., 2007) and for academic success in general (Biemiller & Slonim, 2001; Cunningham & Stanovich, 1997). Direct vocabulary instruction is an essential part of the K-12 curriculum, and there are numerous special programs meant to expedite vocabulary growth in low-income students (e.g., Beck & McKeown, 2007; Goldstein et al., 2017; Gonzalez et al., 2010; Zipoli et al., 2011) and English-language learners (e.g., Carlo et al., 2004; Chung, 2012; Silverman & Hines, 2009). These programs are not without distinguishing features, but common to most are the principles of rich instruction.

When providing rich vocabulary instruction, teachers present words in a variety of meaningful contexts and encourage learners to process those words deeply via comparison, description, and play (Beck et al., 2008; National Reading Panel, National Institute of Child Health and Human Development, 2000). Rich vocabulary instruction incorporates repeated, spaced exposure and extensive practice, factors known to enhance learning in general (Karpicke & Roediger, 2008; Karpicke & Bauernschmidt, 2011) and word learning in particular (Gordon, 2020; Haebig et al., 2019). The targets are typically Tier 2 words, like *introduce* and *compare.* Tier 2 words are termed 'general academic words' in the Common Core Standards. Beck et al. (2013) recommend that teachers focus on Tier 2 words because Tier 1 words, like *happy,* are easily acquired without direct instruction and Tier 3 words, like *isotope,* are so rare as to be of limited functional value. In contrast, Tier 2 words are frequent and applicable across domains.

Rich instruction improves word learning in typical kindergartners (Coyne et al., 2009) and first graders (Maynard et al., 2010) relative to incidental exposure or exposure plus definitions during story reading. Rich instruction has also proven effective at improving the expressive vocabulary of typical kindergarteners from low socioeconomic backgrounds (Beck & McKeown, 2007). But to what extent do rich instructional practices support vocabulary growth in children whose language development is atypical? Vocabulary deficits are a common symptom of atypical language development across populations, including children with Intellectual Disabilities, Autism Spectrum Disorder, and Language Disorder. Without specialized intervention, these children are unlikely to meet the academic standards for vocabulary learning that are expected of them at school. However, it is not clear whether these children also benefit from rich vocabulary instruction. In this descriptive multiple case study, we explore the utility of rich instruction for children with atypical language development by following seven children with Developmental Language Disorder (DLD) as they receive a vocabulary intervention embedded in the science lessons of a five-week summer science camp.

The vocabulary problems associated with DLD

DLD, also termed speech-language impairment, expressive(-receptive) language disorder, specific language impairment, or language delay, is a highly prevalent and persistent neurodevelopmental disability that limits expressive or receptive language ability and that, as a result, interferes with communication in everyday contexts (Bishop et al., 2016). As a group, people with DLD have difficulties learning new words and, therefore, their vocabularies are sparser and shallower than those of their typical peers (McGregor et al., 2002; McGregor et al., 2011; Stothard et al., 1998). Vocabulary deficits among kindergarteners with DLD persist into adolescence (McGregor et al., 2013).

Although the nature of the vocabulary deficits associated with DLD will vary from child to child, there are some common patterns. Both receptive and expressive vocabulary scores may be low, but expressive vocabulary tests tend to be more sensitive for identifying the deficit (McGregor et al., 2012). The problem is both one of breadth—they do not know as many words as other children—as well as depth—they do not have rich representations of known words. Words that are rare or abstract are more problematic than words that are frequent and concrete (McGregor et al., 2012). Verbs may be more problematic than nouns (Windfuhr et al., 2002), and word form may be more problematic than word meaning (McGregor, Arbisi-Kelm et al., 2017). People with DLD need even more exposures to words (i.e., a higher dosage) than their peers to demonstrate comparable learning. The absolute dosage likely varies with age, severity, the number of words to be learned, and whether learning is measured on receptive or expressive tasks (Gray, 2003, 2004; McGregor et al., 2020).

The word learning problems of people with DLD are hypothesized to arise from deficits in short term memory and other aspects of executive function. Executive functions are a set of processes that control and coordinate attention and memory so that we can concentrate, think flexibly, and resist ill-advised impulses (Diamond & Lee, 2011). People with DLD have documented weaknesses in multiple aspects of executive function, including verbal working memory (Archibald & Gathercole, 2006; Kapa & Erikson, 2020), verbal short-term memory (Archibald & Gathercole, 2006; McGregor, Gordon et al., 2017), sustained attention (Ebert & Kohnert, 2011, Kapa & Erikson, 2020; Smolak et al., 2020), and attention shifting and inhibition (Kapa & Erikson, 2020; Pauls & Archibald, 2016). Weiland et al. (2014) examined the relationship between executive function and vocabulary outcomes among a large group of four-year-olds instructed via Opening the World of Learning (OWL; Schickedanz & Dickinson, 2005), a curriculum that includes direct vocabulary instruction. After controlling for vocabulary levels at the beginning of preschool, they found that executive function skills accounted for significant variance in receptive vocabulary at the end of preschool. For children with DLD as well, word learning outcomes are positively correlated with executive function as measured on tasks of inhibition, verbal short-term memory, and verbal working memory (Alt & Plante, 2006; Kapa & Erikson, 2020).

Vocabulary intervention for children with DLD

Although vocabulary is a frequent focus of language intervention for children with DLD (Steele & Mills, 2011), speech-language pathologists (SLPs) tend not to apply all principles of rich vocabulary instruction during these interventions. They target Tier 1 words more often (Justice et al., 2014) or as often (Steele, 2020) as Tier 2 words. In treatment sessions with kindergarteners and first-graders with DLD, SLPs were observed to use some aspects of rich instruction like providing child-friendly definitions and pointing out synonyms, but not others, like recontextualization (do you ever *assist* your mom with anything at home) or demonstration (show me *searching*; what would it look like?) (Justice, et al., 2014). The extent to which they should use more complete rich instructional practices depends on the evidence behind these practices. We turn there now.

Marulis and Neuman (2010) conducted a meta-analysis of 67 vocabulary intervention studies, together including 5,929 preschool and kindergarten children. Most of these involved shared storybook reading although some incorporated other activities as well. Overall, they found a large positive effect of rich instruction when learning outcomes were tested immediately after the intervention and a moderate positive effect when retention was considered. Interventions that provided explicit, rich instruction were more efficacious than those that involved implicit exposures only. So, for example, engaging children in discussion during the book readings worked better than simply reading the book aloud to them.

Critical to our purpose, Marulis and Neuman (2010) examined whether children deemed ‘at-risk’ benefitted from the interventions. These were children who experienced one or more of four risk factors: low socioeconomic status (SES), English language learner status, low academic achievement, or special needs as determined by receipt of an Individualized Education Program or IEP. They found that children who are at-risk reaped equal benefit from direct vocabulary instruction as their low-risk peers. However, children who were from low-SES families AND had at least one additional risk factor did not fare as well. This conclusion was confirmed by Marulis and Newman (2013) in a subsequent meta-analysis that focused exclusively on at risk-children and also by Goldstein et al. (2017), who found large positive effects of an automated rich vocabulary instruction program for low-income children but smaller effects for children who were low-income plus had an IEP, the majority of whom had DLD.

Whereas Marulis and Neuman (2010) grouped children with DLD into a broad category of at-risk learners, others have examined them separately from children with other risks. S. Gillam et al. (2014) tested the effect of rich vocabulary instruction embedded in the discussion of wordless picture books. In that study, an SLP targeted eight Tier 2 words in three 30-minute periods per week for six weeks within the context of a first-grade classroom. The SLP used strategies such as visual supports and elicited production of the words and their definitions within the meaningful context of the books. All children in the classroom received the intervention and, as a group, they made more gains than children in classrooms who received their usual curriculum. However, the effect for those who scored at or below the 25th percentile on the *Test of Narrative Language* (Gillam & Pearson, 2004), a proxy for DLD, was more modest than the effect for more able language learners.

To boost word learning among children with language deficits, Loftus et al. (2010) taught eight words in four daily lessons for two weeks in a kindergarten classroom setting and then supplemented that instruction with small-group intensive practice on four of the eight words for the 20 kindergarteners with the lowest receptive vocabulary levels. As a group, all children benefitted from rich instruction in the classroom, and the 20 with vocabulary deficits made similar gains to their peers on the four words that received supplemental instruction. In other words, the supplemental experiences served to close the learning gap for those four words.

The reasons for the success of the supplemental intervention may be numerous. Within a smaller group, the children may have received scaffolds and prompts that were more specific to their individual needs. Moreover, the supplemental experience provided extra opportunities to hear, use, and engage with the target words (Kong, 2013) and thus establish a stronger memory trace. McGregor et al. (2020) found that adults with DLD need more opportunities, that is a higher dosage than their peers, to master words at comparable levels. Once given this dose, they retained the mastered words equally well.

Storkel and colleagues have explored the dosage issue. Storkel, Voelmle et al. (2017) provided rich instruction to kindergarteners with DLD in shared storybook readings and manipulated the dosage received. The children learned fewer words with a dosage of 12 exposures than 24, and fewer with a dosage of 24 than 36. However, increasing the dosage to 48 did not improve learning. They concluded that 36 exposures, which equates to three times the intensity required by typically developing children undergoing the same instruction (Justice et al., 2005), was adequate.

Storkel et al. (2019), asked whether there was a best way to package these exposures. Do six exposures to each word in each of six different sessions work better than four exposures in nine sessions? How about nine exposures in four sessions? In all three conditions, the kindergarteners learned to define 9 or 10 of the 30 target words, on average. This is a hopeful finding as it suggests that SLPs have some flexibility in how they schedule doses over sessions provided that an adequate cumulative dose is delivered. However, there were two caveats. Not all children learned well. After 36 exposures to each word, the strongest learner could define 21 but the weakest learner could define none. Performance varied with severity as measured by the Core Language composite scores on the *Clinical Evaluation of Language Fundamentals* (Semel et al., 2003). In other words, the children who likely had the greatest needs learned the least. Moreover, two weeks after the treatment was withdrawn, the children defined only 3 of the 30 words, on average (range = 0 to 14), down from the original average number of 9 or 10 learned words. In discussing potential causes of forgetting, Storkel et al. (2019) note that their intervention, while semantically rich, did not require active engagement as recommended for rich vocabulary instruction (Beck et al., 2008, 2013). Also, their approach did not involve revisiting trained words, a practice that provides review and opportunities for re-encoding. It could be that the nonresponders needed periodic reviews or stronger initial learning via active engagement to stave off forgetting.

In summary, rich vocabulary instruction is demonstrated to be effective for typical learners. It also benefits learners with DLD but, in some cases, to a lesser extent. Children with severe DLD or those who have DLD accompanied by the disadvantages of poverty may demonstrate minimal gains. That said, the evidence base is incomplete because not all studies have employed a wide variety of rich instruction strategies, active engagement, a therapeutically high dose, or review.

Current study

In this study, we describe the experience of seven kindergartners with DLD who received rich vocabulary instruction that involved direct teaching of Tier 2 science-relevant vocabulary. The instruction was designed to offer a therapeutically high dose, motivate active engagement, and provide opportunities for review. The intervention was embedded in a hands-on, inquiry-based science curriculum. The science curriculum itself was language-rich and it provided an authentic context in which to implement rich instruction. Moreover, it enabled a parallel to school-based language interventions that often link to the academic curriculum and take place in the classroom via collaboration or consultation or in small group pull-outs (Steele, 2020).

In comparing and contrasting these seven cases, we not only considered their vocabulary outcomes but also explored the characteristics of the children, the clinicians' delivery of rich instruction, and the vocabulary words themselves to highlight potential reasons for differences in outcomes. Via case-study description, we aimed to support translation from research to clinic or classroom by sharing treatment successes as well as failures for the consideration of the SLPs who provide these interventions to children.

Method

Overall project

The vocabulary intervention was one of three language interventions that took place during a summer science camp. The overall project was structured as a randomized controlled trial, funded by the National Science Foundation, and approved by the IRB of the University of Delaware.

Participants. Participants were recruited via area special education districts, school districts, private SLPs, preschools, partnerships with community agencies, and magazine ads. We included only children who had not yet attended first grade, this to ensure that the science lessons, derived from a grade 1 curriculum, would be unfamiliar to all children. Enrollees had to be primarily English speaking, which was defined as less than 20% exposure to a language other than English via parent report. To ensure that they could participate in the language of the lessons, enrollees had to be able to produce simple sentences. Because we were interested in children with DLD not associated with other biomedical conditions we ruled out hearing loss, intellectual disability, and other potentially causal conditions by enrolling only those children who passed a pure tone audiometric screening administered according to the standards of the American Speech-Language-Hearing Association (1997); earned a percentile higher than 7 on the Differential Abilities Scale-II, matrices subtest (DAS-II, Elliot, 2007); scored below 15 on the Social Communication Questionnaire (SCQ; Rutter et al., 2003); and had negative histories of Autism Spectrum Disorder and genetic syndromes via parent report. The SCQ is a parent-report instrument that summarizes problems in social relating, communication, and range of interests. Children who score above 15 may be on the autism spectrum and should receive a full evaluation to determine a correct diagnosis. These children were referred to other professionals and not enrolled in camp, regardless of the outcome of that diagnosis.

To verify potential DLD we took several steps. First, children had to be receiving speech-language services at school or, for parent and community referrals without a history of services, had to score in the range of clinical concern on a nonword repetition task (Dollaghan & Campbell, 1998). Poor nonword repetition is a clinical marker of DLD (Archibald & Gathercole, 2006). We used age-specific cutoffs demonstrated to maximize sensitivity to the identification of DLD. For four- and five-year-olds, the cutoff was 67% of phonemes correct (Deevy et al., 2010). For six-year-olds, it was 75% (Dollaghan & Campbell, 1998).

Next, we administered the *Diagnostic Evaluation of Language Variation-Screening Test* (DELV-ST, Seymour et al., 2003b) to determine whether the child spoke Mainstream American English. If so, the child took the *Structured Photographic Expressive Language Test-3* (SPELT-3, Dawson et al., 2003). The SPELT-3 measures children's production of grammatical morphology and syntax. We enrolled only those children who scored below a standard score of 95 (the 37th percentile) on the SPELT-3, this being the cut-off that maximizes the sensitivity and specificity of identification of DLD (Perona et al., 2005).

If the child did not speak Mainstream American English, we administered the syntax subtest of the *Diagnostic Evaluation of Language Variation-Norm Referenced* (DELV-NR, Seymour et al., 2003a). This subtest measures children's knowledge of wh-questions, passives, and articles, syntactic constructions that are universally diagnostic across dialects of English. We applied a traditional clinical cut-off, the 10th percentile, because we were not aware of any research determining the cut-off that maximizes sensitivity and specificity of the identification of DLD on this test.

Three of the children in the vocabulary intervention were from low-SES families, as determined by their need for State-provided purchase of care vouchers for food. These families were highly motivated to enroll their children in the science camp and its associated language intervention; however, they experienced a number of challenges during the period of the summer camp that at times limited the children's participation and, perhaps, the benefit of the experience. We cannot say for sure that these challenges were the direct or indirect result of poverty, nor can we say with certainty that the more affluent families who participated were not experiencing a comparable number or similar type of challenges, although they did not mention any challenges to us. Nevertheless, listing these challenges may provide some insights into the repeated finding that children with DLD who are from low income families tend to gain limited benefits from speech-language interventions. In other words, it may help us to understand why low SES is a useful proxy for prognosis. Across the three intervention arms, families who required food vouchers communicated the following to us while we took case histories and organized camp logistics or talked to them during drop-off and pick-up times:

> Father lost his job; Father enrolled in a drug and alcohol rehabilitation program; Family moved for the second time in six months; Family home had no electricity; Mother with no work history was seeking a job; Family lacked transportation; Family received purchase of care vouchers because they could not afford childcare; Family received State-provided visiting nurse support because of at-risk newborn; Child broke his/her arm; Caregivers had highly unpredictable work schedules; Parent broke a leg and therefore cannot drive child to camp; Parent sent to jail. More than three primary caregivers coordinating care. Parent recently released from jail. Child adopted out of the foster system.

When we present the individual cases below, we will indicate the number but not the type of challenges that the individual faced to protect anonymity.

Materials. The curriculum was an adaptation of the inquiry-based Full Option Science System-Next Generation Edition for grade one (FOSS, Lawrence Hall of Science, 2015). The science lessons covered topics such as plant growth, light, sound, and air. The sequence of activities and supports provided throughout each session were based on those recommended in the FOSS Investigations Guide. FOSS was selected because there is evidence for its effectiveness (Nowicki et al., 2013; Young & Lee, 2005) and because the hands-on, yet structured, nature of the FOSS curriculum is consistent with best practices for teaching science to children with learning challenges (Brigham et al., 2011; Therrien et al., 2011). The science activities provided a meaningful, engaging context in which to model and elicit language targets. The teacher's manual included scripts that were followed faithfully with the exception that we included journaling and book reading in every session whereas these were more occasional in the FOSS lessons.

Approach. Participants attended a five-week summer science camp for 3.5 hours per day, five days per week. The first four of the five weeks included a language intervention. The participants were randomly assigned to one of three intervention arms: vocabulary, grammar, or phonological awareness. The three intervention arms occurred in three separate classrooms, each staffed by three graduate student clinicians and one master clinician. Each student clinician was enrolled in a graduate program in speech-language pathology. The student clinicians had completed at least one course on language development and two more on childhood language disorders. Students completed an application to participate that included assessing their prior knowledge of science, grammar, vocabulary, and phonological awareness. The master clinicians were certified SLPs who had either master's or doctoral degrees and at least three years of experience with childhood populations. An additional 12 undergraduate students assisted by preparing science materials, engaging in fidelity checks, providing before/aftercare, transcribing responses, or supporting other data analysis activities.

The daily schedule began with a morning greeting song with all campers, followed by a 40-minute science + language intervention session, snack and outdoor play, and finally, a second 40-minute science + language intervention session. During each session, we followed the four steps of the FOSS curriculum: 1) Prediction: background information was discussed, pre-existing knowledge and beliefs were identified, and participants were guided to generate predictions about the topic; 2) Experiment: participants were guided through the experiment/observation; 3) Journal/Reflection: participants completed a science journal by drawing, adding materials (e.g. glue + leaves) and dictation to the adult; and 4) Text-based research: the group engaged in the dialogic reading of an expository text linked to current content. Steps 1 to 3 were typically accomplished in small groups of two to five, led by a student clinician. After step 3, all children and clinicians in the intervention classroom (7 children, three student clinicians and one master clinician) came together for circle-time, where they shared their journals and participated in the book reading activity. Notice that these steps are already language-rich; they involve explanation, conversation, dialogic reading, and dictation. We did not strip this rich language from the science lessons; rather, for each intervention, we integrated additional, targeted supports according to the intervention condition.

Vocabulary intervention

The intervention targets were 20 Tier-2 words from *Vocabulary for the New Science Standards* (Marzano Research, 2015). Each word was relevant to one or more of eight scientific practices that are appropriate for children in grades K to 2: questioning, modeling, investigating, gathering data, computing, explaining, arguing, and communicating (National Research Council, 2012).

The therapeutic approach was inspired by the book *Robust Vocabulary Instruction* by Beck et al. (2013). It involved modeling the new words via four strategies: introducing words along with child-friendly definitions as they are encountered (e.g., "A *hypothesis* is a guess about what will happen."); contextualizing or linking providing examples of the word within the meaningful context of the science lesson (e.g., "We're making a *model* habitat."); recontextualizing or helping the child to associate new words with their own experience (e.g., ″Did you ever have to *search* for something you lost?"); and engaging in playful routines (e.g., "Every time you see a *scientist* in this story clap your hands!"). Labeling, verbally or in writing, target words in the child's journal was included as a special case of contextualizing. The clinician also elicited productions of the words by asking questions and following with the necessary cues to support the child's answer.

The student clinicians were trained to deliver rich vocabulary instruction during 'minicamp,' a four-day science camp for children who did not qualify for the intervention study but who, nevertheless wanted a camp experience. Prior to minicamp, the student clinicians in the vocabulary arm read Beck, et al. (2008) and McGregor and Duff (2015) and submitted any questions they had for discussion with the first author and the master clinician. They were given the guidelines that appear in Supplemental Materials-1 (McGregor, 2020) and these were modeled for them. The students practiced implementing these within the context of science instruction during minicamp. Any student who was not providing the minimum number of models and elicitations or using the four modeling strategies and the elicitation cueing hierarchy was given individual coaching by the master clinician. During the actual intervention, the master clinician scaffolded the student clinicians heavily in the beginning and gradually lessened the scaffolds as the clinicians grew more able. A student observer tracked the fidelity of 1/3 of the science + vocabulary lessons live by rotating her observations across the student clinicians. She used the FOSS scripts and the vocabulary strategies list to structure the observation.

Targeting a smaller set of words at a time is more efficient than training larger sets (Beck et al., 2008; Johnson et al., 1987); therefore, each science lesson featured two of the target words. We used a cyclic goal-attack strategy that mirrors typical practice in U.S. schools (Steele, 2020). Specifically, each word was targeted in four lessons spaced over four weeks so that the children's knowledge of the words could be refreshed and enhanced periodically. Two words, *diagram* and *material*, were targeted in three lessons only due to time constraints. Each word was to be modeled at least 3 times during the hands-on science lesson and at least once during the shared book reading. Each was to be elicited at least once per lesson. Therefore, the cumulative minimal dosage per word would be 20 ([4 models + 1 elicitation] x 4 lessons). This minimum was derived by considering that rich vocabulary instruction in a dose of five exposures yields large effects for typical children (Beck & McKeown, 2007) and reasoning that a dosage that was four times larger would suffice for children with DLD. However, some of the words appeared in the scripted FOSS lessons themselves and, to remain true to the *Robust Vocabulary Instruction* guidelines, the clinician was free to provide additional models and elicitations as the meaningful context and engagement of the children allowed. Thus, in practice, the doses were larger than 20. In fact, in circle-time alone, where the minimal cumulative exposure was set at four, the actual minimum received for any given word was 6 and the maximum was 38.

Outcome measurement. We measured the children's science learning with pre- and post-intervention probes derived from the FOSS unit tests that tapped knowledge of trained content. Unit tests were adapted to avoid words taught during the vocabulary intervention itself and to minimize the need for verbal responses. Example questions include:

- Which one has the highest pitch? Which one has the lowest pitch? (when showing the child 3 bars of varying lengths)

- Find the *root.* Find the *stem*. Find the *leaves*. (on a picture of a plant).

We measured vocabulary outcomes in two ways. First, because the journal sharing and book reading activities were always video-taped, we could assess the children's spontaneous productions of the targets within a meaningful context. Second, we devised a 3-alternative-forced-choice (3AFC) probe to determine pre-intervention knowledge of the words, post-intervention learning, and long-term retention of any learned words. The long-term probe was administered six months after camp had ended. Only three of the seven children in the vocabulary intervention were available for long-term follow-up. On the probe, a child heard a sentence that contained a target word (e.g., Show me *collecting* rocks) and picked an answer from among three possible pictures (e.g., examining rocks, collecting rocks, and climbing rocks). We compared the children's performance to chance (33%) to make inferences about whether words were or were not learned.

We considered the children's vocabulary outcomes in light of their language and executive function abilities. To profile their comprehension and learning processes in the domains of syntax and vocabulary, we administered the *Quick Interactive Language Screener* (QUILS, Golinkoff et al., 2017) to each child before camp began. Although we did not use this test as part of our enrollment criteria, it is useful to note that a score in any domain—vocabulary, syntax, or process (the ability to learn new language items)— that falls below the 25th percentile is a concern and cause for more detailed assessment in that area. The *Peabody Picture Vocabulary Test*-fourth edition (PPVT-4, Dunn & Dunn, 2007) and the *Expressive Vocabulary Test*-second edition (EVT-2, Williams, 2014) are companion tools for evaluating age-appropriate knowledge of English vocabulary words. Although we used these tests for description only, for the purpose of interpretation, note that a cut-off of 1 standard deviation below the mean (the 16th percentile) maximizes sensitivity to DLD (Pearson, 2018).

Prior to camp, the parents of the participants completed the *Behavior Rating Inventory of Executive Function-Preschool* (BRIEF-P, Gioia et al., 2003), a survey that rates the child's ability to **inhibit** or control impulses; **shift** or transition from one task to the next and problem-solve flexibly; **control** emotional responses; **initiate** or begin activities and generate ideas; use **working memory** to hold ideas in mind to aid in task completion; **plan and organize** steps for reaching goals; and **monitor** or assess their own performance. Ratings in these seven domains are combined to determine a global executive composite. The mean T score is 50 (SD = 10) and scores higher than T = 65 (percentile = 93) are considered clinically significant. At the end of camp, when the master clinician had observed each child for five weeks, she completed the BRIEF for all campers as well.

Results

The children

The children who received the vocabulary intervention were one girl and six boys. As a group, these children learned from the science curriculum as well as the vocabulary intervention. They averaged 29% accuracy on the science pretest (SD = 9%) and 42% accuracy on the posttest (SD = 11%), $t = 5.15$, $p = .004$. The minimum gain was 7.5% accuracy and the maximum gain was 22.5%.

The children averaged 36% accuracy on the vocabulary pretest (SD = 12%, min-max = 20% to 50%). This performance was not better than chance, $t = .61$, $p = .56$. After the intervention, they averaged 53% accuracy (SD = 15%, min-max = 40% to 80%), a value that exceeded chance level performance, $t = 3.38$, $p = .02$. The difference between vocabulary pretest and posttest performance was significant, $t = 4.34$, $p = .004$.

In the following, we describe each child's presentation and response to the intervention. We have ordered the cases by the amount of progress they made in response to the vocabulary intervention, from most to least. The standardized test scores of the participants appear in Table 1; note that all names are pseudonyms.

Table 1.

The children's demographics and standardized test scores expressed as percentiles.

	Elvis	Porter	Sanjay	Adam	Luke	Brianna	Tyrone
Sex	boy	boy	boy	boy	boy	girl	boy
Age	6.64	4.10	4.05	5.54	4.29	4.40	4.77
Race, ethnicity	W, NH	W, NH	A, NH	W, NH	W, NR	B, NH	B,NH
Parent Education (years)	14, 10	18, 18	12, 14	18, 17	14, 16	14	12, 12
Attendance (%days)	72.5%	100%	100%	90%	90%	80%	55%
SPELT-3			13*				
DELV-NR	5*	5*		3*	4*	1*	9*
DAS-matrices	14	42	14	34	79	14	34
QUILS overall	65.8	21.3*	13.3*	6.2*	10.2*	17.0*	21.3*
QUILS receptive syntax	91.6	13.9*	6.5*	9.2*	25.3	13.0*	13.0*
QUILS receptive vocabulary	58.0	15.6*	22.1*	5.3*	22.1*	27.3	22.1*
QUILS processing	26.7	27.9	14.9*	1.5*	1.3*	14.9*	20.8*
PPVT-4	37	37	8*	3*	45	32	7*
EVT-2		30	34	14*	73	68	9*
BRIEF Global Executive**	42,-	84,98*	54,75	99.9*,95*	>99.9*,84	96*,-	>99.9*, >99.9*

*score indicates potential clinical concern, **In some cases BRIEF scores were based on clinician judgement and in other cases on both clinician and parent; the clinician's judgement is listed first, followed by the parent's (e.g., clinician score, parent score).
W = White; NH = not Hispanic; B = Black or African American; A = Asian; NR = NR; SPELT-3 = Structured Photographic Language Test-3; DELV-NR = Diagnostic Evaluation of Language Variation-Norm Referenced; DAS = Differential Abilities Scale; QUILS = Quick Interactive Language Screening; PPVT-4 = Peabody Picture Vocabulary Test-4; EVT = Expressive Vocabulary Test-2; BRIEF = Behavior Rating Inventory of Executive Function.

Good responders

Elvis demonstrated the best response to intervention of any camper in the vocabulary group. On the vocabulary pretest, he answered 50% of the items correctly. On the posttest, he answered those same 10 of 20 items correctly, plus an additional six, for an overall accuracy of 80% (see Table 2). Elvis had the highest science knowledge of any camper at pretest (46% accuracy on the science pretest) and at posttest (60% accuracy on the science posttest).

Table 2
Pretest to posttest changes by word and case.

	Elvis		Porter		Sanjay		Adam		Luke		Brianna	
	pretest	posttest	pretest	posttest	pretest	posttest	pretest	posttest	pretest	posttest	pretest	posttest
Collect	+	+	+	+	+	+	+	+	-	**+**	+	+
Compare	-	-	-	**+**	-	-	-	-	-	-	+	+
Data	+	+	-	-	-	**+**	-	**+**	-	**+**	+	**-**
Design	+	+	+	+	-	**+**	-	**+**	-	-	+	+
Diagram	-	**+**	-	-	-	-	-	-	-	-	-	-
Evidence	-	**+**	-	-	-	-	-	-	-	-	-	**+**
Explanation	-	-	-	-	-	-	-	**+**	+	**-**	+	**-**
First-hand	-	**+**	+	+	-	**+**	+	**-**	-	-	-	-
Hypothesis	-	-	+	**-**	+	**-**	-	**+**	+	**-**	-	-
Material	+	+	+	+	+	+	+	**-**	+	+	-	**+**
Model	+	+	-	**+**	-	-	-	-	+	+	-	**+**
Multiple	-	**+**	+	+	-	**+**	-	**+**	+	+	+	**-**
Pattern	-	-	-	-	-	-	-	-	-	-	+	+
Plan	+	+	+	**-**	+	+	-	**+**	-	-	+	+
Problem	-	**+**	-	**+**	-	-	-	-	-	**+**	+	+
Scientist	+	+	-	**+**	-	-	-	-	-	-	-	-
Search	+	+	+	+	-	**+**	-	-	-	**+**	-	-
Solve	+	+	-	**+**	+	+	+	**-**	+	+	+	+
Study	+	+	-	-	-	-	-	**+**	-	-	-	-
Test	-	**+**	-	**+**	-	-	-	-	-	**+**	-	-

Note: Because Tyrone did not complete the posttest, we removed his pretest performance here to enable comparison and determination of gains. Bold font = change in accuracy from pretest to posttest.

Details of Elvis's birth and health history were not provided on the case history form, but his parents did report a sibling who also had DLD. Although we did not test for this directly, we noted that he had a severe speech sound disorder. He received speech therapy at a private clinic and also at school where his IEP was exclusively focused on phonology and repair strategies given his poor intelligibility.

Elvis was one of the three campers from a low-SES background and we came to know about five of the stressors listed above. Some of these stressors caused Elvis to miss more than 25% of all camp days and to be more than an hour late to camp on other occasions. Thus, he received a lower dose of the vocabulary intervention than other campers. Nevertheless, Elvis had some notable advantages over the other campers. At 6.5 years, he was the oldest. He had the mildest presentation of DLD with a clinical concern in syntax only, as indicated by the DELV-NR. He had intact executive function and he remained attentive and engaged at a fair level throughout camp.

Porter correctly identified 40% of the vocabulary items at pretest and he improved to 60% at posttest. He was observed to use the word *design* spontaneously during circle-time, the only child observed to utter any of the words without prompting. His science pretest score was 27.5% and his posttest was 35%. Porter had no significant birth or health history. Nevertheless, he had been slow to meet speech-language, fine-motor, and social-emotional developmental milestones and he received specialist intervention in all three of those domains. His parents reported that he avoided talking to other children during play and at school and that he sometimes cried when communicative interactions failed, a pattern of behavior also observed at camp. Porter's performance on the test battery indicated clinical concerns in syntax (DELV-NR, QUILS) and vocabulary (QUILS). However, vocabulary, as measured by the PPVT-4 and the EVT-2 was a relative strength. Porter's parents reported concerns with executive function in the home environment, but we did not see evidence of executive function deficits at camp. Porter's parents reported that he loved science, and this interest was reflected in his consistently good attention to and engagement in the science lessons. He was present for all camp days, so he received a full dose of the intervention. His strengths in vocabulary and executive function and his engagement in the camp and its associated intervention likely contributed to his good outcome.

Elvis and Porter brought unique learning strengths to the intervention. Elvis had stronger language ability and better prior knowledge of science, but Porter had heightened interest in science activities. Elvis's family reported significant challenges during camp but Porter's family reported none. Although, in our opinion, both had intact executive function skills for their age, Elvis was older by more than two years, thus his absolute executive function was higher. Finally, Elvis missed camp frequently, but Porter never missed, giving Porter an advantageous dosage.

Fair responders

Sanjay and Adam demonstrated a 20% improvement on the vocabulary probe after the intervention. Although Porter, too, had demonstrated a 20% improvement, we judge his response to intervention to be more robust because his final score was well above chance and he learned at least one of the words well enough to use it spontaneously. In contrast, Sanjay and Adam hovered on either side of chance before and after the intervention and never used any of the words spontaneously.

Sanjay was a boy with an unremarkable birth history. His parents were bilingual, but they reported that he was exposed to English only. Sanjay was reported to have more than four ear infections per year. His speech-language development had been slow, and he had been diagnosed with DLD for which he received speech-language intervention. His siblings also had speech-language impairments. Sanjay's scores indicated concerns in syntax (SPELT-3), receptive vocabulary (PPVT-4), and overall language learning ability (QUILS). Nevertheless, expressive vocabulary was a strength and he had age-appropriate executive function skills. The summer camp was Sanjay's first experience with a structured peer group, having never been enrolled in childcare or preschool programs. He adapted well and was judged to have a good level of engagement and attention during camp. His science pretest score was 26% and his posttest score was 46%. On the vocabulary probe he was 25% accurate at pretest, 45% accurate at posttest, and 60% accurate on the delayed posttest administered six months after camp had ended. He began preschool between camp and the delayed posttest.

Adam was a boy with no significant birth or health history; however, he did not combine words until he was 3 or 4 years old. He received speech-language intervention at his school where his IEP goals focused on academic concepts, language, and social skills. Adam's parents said that he avoided talking to adults and that he preferred numbers to reading. His test scores indicated severe DLD affecting both syntax and vocabulary; in fact, his scores were the lowest of all the children in the vocabulary intervention. His camp clinicians also observed some social-pragmatic weaknesses. Adam demonstrated clinically significant executive function concerns on the BRIEF, as reported by both the parent and the clinician, in the areas of inhibition, attention shifting, working memory, and planning, with working memory being the weakest. Adam's engagement during camp was judged to be fair and he demonstrated modest gains in science, from a pretest score of 26% accuracy to a posttest score of 36% accuracy. On the vocabulary pretest, he identified 20% of the items correctly. On the posttest, he was 40% correct. Adam was seen for a long-term follow-up six months after camp and, at that time, his score had increased to 50%. Between camp and the follow-up, he began full-day kindergarten and continued to receive speech-language services.

Although Sanjay and Adam both improved from pretest to posttest, it is difficult to evaluate their learning because their scores on both pre- and post-tests hovered on one side or the other of 33%, the chance level of responding. Therefore, we examined the stability of their answers for each item from pre- to post-test, thinking that true chance would result in high variability in both directions (items that were incorrect at pretest might be correct at posttest but also, items that were correct at pretest might be incorrect at posttest). In contrast, if their responses reflected their prior knowledge and learning during the intervention, they would remain stable on some items with others moving from incorrect to correct.

Sanjay demonstrated stable responding for 14 words (Table 2). He correctly identified *hypothesis* at pretest but not at posttest, but his changes for five other words, *search, multiple, design, data*, and *firsthand*, moved in a positive direction. Adam demonstrated stable performance on 10 words. He correctly identified three items at pretest, *solve, first-hand*, and *material*, but missed these at posttest. In contrast, he changed from incorrect to correct on the remaining seven items: *data, design, explanation, hypothesis, multiple, plan*, and *study*. We conclude that Sanjay and Adam demonstrated modest learning. Their subsequent improvement from the posttest to the delayed posttest provides converging evidence. Their knowledge of the new words they had learned may have been reinforced at school and their memory for those words may have consolidated over time.

Sanjay had stronger language and executive function skills than Adam, but Adam was a year older than Sanjay. As with Elvis and Porter, they had similar outcomes despite bringing different advantages and learning strengths to the intervention.

Minimal responder

Luke was a boy who had been born full-term and with no complications despite placenta previa. His parents reported good health and typical development except for concerns about his speech and language abilities. He received speech-language intervention at school with goals focused on speech, language, and social behavior. His performance on the DELV-NR syntax and the QUILS processing was low but his vocabulary, as measured by the PPVT-4 and the EVT-2 appeared to be a strength, in fact, they were the strongest of the group. As with Adam, his camp clinicians judged his social-pragmatic skills as weak. According to the clinicians, Luke exhibited weaknesses in executive function during camp. His executive composite score reflected difficulties in the areas of inhibition, shifting, working memory, and planning, with working memory being the weakest. However, his parents did not report significant concerns with his executive function at home. His attention and engagement during camp were judged as fair. He demonstrated the largest gains in science knowledge of any camper, moving from a science pretest accuracy of 26% to a posttest accuracy of 49%. His pretest accuracy on the vocabulary probe was 30% and his posttest accuracy was 45%.

Luke demonstrated stable performance on 13 words. He flipped from correct to incorrect on two words, *explanation* and *hypothesis*, but he flipped from incorrect to correct on five: *collect, data*, *problem, search,* and *test*. We conclude that Luke demonstrated minimal learning. We might expect that Luke would have been a stronger word learner than Sanjay and Adam because their profiles were similar except that Luke had notably better vocabulary and nonverbal cognition. However, Luke had weaker executive function skills than Sanjay and he was a year younger than Adam. These differences may have limited his relative success.

Nonresponder

Brianna was a girl from a low-income family in good health and with no developmental concerns other than speech and language. Brianna received speech-language services at her community preschool. Her performance on the test battery indicated clinical concerns in overall language knowledge and learning ability (QUILS), syntax (DELV-NR), and executive function (BRIEF). Particular areas of concern on the BRIEF were inhibition, working memory, and planning. Although we did not formally measure oral-motor function, we observed poor oral-motor control and drooling. Vocabulary, as measured by the QUILS, PPVT-4 and the EVT-2 were relative strengths. Brianna's family experienced four stressors from the list above during the camp period. Perhaps as a result, Brianna missed 20% of the camp sessions. When present, her attention was judged as consistently poor. She demonstrated minimal gains in science, moving from 21% accuracy on the science pretest to 29% on the posttest. On the vocabulary probe, she was 50% accurate at pretest, 50% accurate at posttest, and 45% accurate at the long-term follow up six months later. She was enrolled in a full-day special education preschool between camp and the delayed posttest.

Brianna, our weakest responder, and Elvis, our best, are interesting to compare. They both presented with DLD and disadvantages associated with poverty. They had similar scores on the DELV, DAS, and PPVT-4 and identical scores on the vocabulary pretest. Their families experienced a number of challenges that resulted in missed intervention sessions and thus, a lower dose than planned. Nevertheless, Elvis flourished and Brianna floundered. Elvis was two years older than Brianna; he had stronger executive function and overall language ability, and he had stronger prior knowledge of science upon enrollment was likely; therefore, he was more ready and able to learn. He was also more engaged in the lessons than Brianna. The age, executive function, language level, science knowledge and engagement likely interacted in ways that helped Elvis to ‘spiral up’ but led Brianna to receive minimal benefit from the intervention.

Attrition

Tyrone was a boy from a low-income family. During camp, his family experienced nine of the stressors described above. His parents did not complete a case-history form but they did report that he received speech-language intervention at school. Tyrone presented with language deficits in all domains according to his standardized test profile, and severe executive function deficits in all domains, according to both the parent’s and clinician’s ratings on the BRIEF. Although Tyrone’s parents did not report a diagnosis of Attention Deficit Hyperactivity Disorder, his scores on the BRIEF are indicative. We did not test for ADHD directly. Tyrone spoke frequently during circle-time but he was often unintelligible and, thus, the clinicians had difficulty responding contingently to him. From the onset of the intervention, Tyrone demonstrated poor attention and disruptive and aggressive behavior toward the staff and the other children. Even with maximal assistance, one-on-one treatment, and significant coaching from the master clinician, he had difficulty participating for more than five minutes at a time. Eventually, these behaviors became so consistent that we took them as his way of withdrawing his assent; therefore, he left the summer camp after 2.5 weeks. Tyrone was 22% accurate on the science pretest and 35% accurate on the vocabulary pretest. We do not have any posttest data.

The clinicians

The graduate student clinicians were two women and one man. They rotated in their assignments so that each clinician had a chance to work with each child over the course of the intervention. This approach also allowed us to respond flexibly to child absences and behavior. In some cases, a clinician worked one-one-one with Tyrone to provide maximum support, and, in other cases, reorganizing groups helped to increase the children’s motivation. The clinicians' fidelity to the principles of rich instruction was determined by observation of their lessons in real-time. Clinician A was observed 13 times, Clinicians B and C were observed 12 times. Each clinician met the minimum dosage of three models and one elicitation for every target word per lesson. In fact, each clinician surpassed the required dosage, provided at least one child-friendly definition per word, and used at least three different instructional strategies per word as directed (Table 3).

Table 3.

The clinicians use of rich vocabulary instruction strategies.

Models, Elicitations, Strategies		Clinician A (N = 13)	Clinician B (N = 12)	Clinician C (N = 12)
Total models	Mean	24.73	26.67	25.5
	SD	(9.18)	(11.31)	(11.17)
	Min to Max	14 - 44	12 - 56	12 - 54
Total elicitations	Mean	5.00	3.86	2.96
	SD	(2.43)	(3.19)	(1.76)
	Min to Max	1 - 9	1 - 12	1 - 7
Provide child-friendly definition	Mean	2.23	1.71	2.88
	SD	(2.92)	(1.12)	(3.66)
	Min to Max	0 - 14	0 - 4	0 - 14
Contextualize	Mean	16.46	20.08	20.00
	SD	(6.54)	(9.06)	(8.11)
	Min to Max	0 - 32	7 - 46	2 - 36
Recontextualize	Mean	4.04	2.92	2.17
	SD	(5.02)	(5.27)	(4.02)
	Min to Max	0 - 19	0 - 18	0 - 18
Engage in routine	Mean	0	0	0
	SD	0	0	0
	Min to Max	0	0	0

N = number of sessions observed; SD = Standard Deviation

We are not only interested in the quantity of the dose but the quality. Beck et al. (2008, 2013) stress that teachers should use a variety of strategies to ensure rich instruction, but the clinicians observed here did not use all strategies equally well. All three clinicians provided about two child-friendly definitions for each target word, per session, on average. For some words, Clinician A and Clinician C provided numerous child-friendly definitions, as many as 14 per word. Clinician B, at most, provided four per word.

Contextualizing was used more often than recontextualizing by all three clinicians and this makes sense given that they were also attempting to remain faithful to the teaching scripts for the FOSS curriculum and their fidelity to those scripts was also being monitored. Modeling and eliciting the target words within the context of the science lesson allowed the clinicians to meet both science teaching and vocabulary intervention goals. That said, Clinician A contextualized a bit less and recontextualized a bit more than Clinician B or Clinician C. None of the clinicians was observed to use game-like routines; perhaps this strategy, in particular, felt too distracting from the science lessons themselves. Note that, because the clinicians rotated amongst the children, these differences in their delivery may have contributed to overall outcomes but they are unlikely to account for the stronger or weaker outcomes of any given child.

The words

We reasoned that the learnability of the target words would vary with phonological complexity and dosage. In Table 4 we specify how many children correctly identified each word at pretest and posttest and group the words according to gains so that their complexity and dosage can be easily compared. Words that demonstrated gains had fewer syllables, on average, than words with no gains. There were no gains for the three 4-syllable words in the set. The consonants that English-speaking children tend to acquire last are /ʃ, ʒ, θ, ð, s, z, l, ɹ/ (Shriberg & Kwiatkowski, 1994). Words with gains had fewer of these so-called late-8 consonants than words without gains.

Table 4.

The target words, their characteristics, models, elicitations, and gains.

Target	IPA	AoA	Syll	Late 8	Lessons	Script models	Circle models	Elicited	Children correct at Pretest*	Children correct at Posttest	Gains
explanation	ɛkspləneɪʃən	7.74	4	3	4	3	12	0	2	1	n
hypothesis	haɪpɑθəsɪs	11.89	4	3	4	2	18	1	3	1	n
material	mətɪriəl	8.06	4	2	3	3	12	0	5	5	n
pattern	pætɚn	6.89	2	1	4	2	28	0	1	1	n
plan	plæn	6.78	1	1	4	6	6	0	4	4	n
solve	sɑlv	6.58	1	2	4	2	15	0	5	5	n
MEAN		7.99	2.67	2.00	3.83	3.00	15.17	0.17	3.33	2.50	
SD		1.82	1.37	0.82	0.37	1.41	6.79	0.37	1.49	1.61	
collect	kəlɛkt	5.61	2	1	4	15	10	2	5	6	y
compare	kəmpɛr	8.57	2	1	4	20	20	0	1	2	y
data	deɪtə	10.42	2	0	4	3	17	1	2	4	y
design	dɪzaɪn	10.05	2	1	4	3	26	2	3	4	y
diagram	daɪəgræm	9.63	3	1	3	1	10	0	0	1	y
evidence	ɛvɪdəns	10.58	3	1	4	9	10	1	0	2	y
first-hand	fɝshænd		2	2	4	0	9	0	2	3	y
model	mɑdəl	5.83	2	1	4	15	34	4	2	4	y
multiple	mʌltəpəl	7.79	3	2	4	2	38	8	3	5	y
problem	prɑbləm	6.06	2	2	4	10	15	0	1	4	y
scientist	saɪəntɪst	6.89	3	2	4	12	13	0	1	2	y
search	sɝtʃ	8.72	1	3	4	12	25	1	2	3	y
study	stʌdi	6.58	2	1	4	12	15	0	1	2	y
test	tɛst	6.26	1	1	4	6	10	0	0	3	y
MEAN		7.92	2.14	1.36	3.93	8.57	18.00	1.36	1.64	3.21	
SD		1.84	0.66	0.74	0.27	6.16	9.42	2.24	1.39	1.37	

*Because Tyrone did not complete the posttest, we removed his pretest performance here to enable comparison and determination of gains. IPA = International Phonetic Alphabet; AoA = Age of Acquisition; Syll = Number of Syllables; Late 8 = eight consonants that are the latest acquired by English-speaking children; SD = Standard Deviation

Table 4 includes multiple proxies for dose per word: the number of lessons, the number of appearances in FOSS scripts, and the number of models and elicitations during circle-time. The number of lessons did not seem to matter: Two words were targeted in three lessons only; one had gains and the other did not. Words without gains appeared less often in the FOSS scripts and less often during circle-time than words with gains. Words without gains were elicited from the children less often than words with gains. Of course, elicitations are both a means and reflection of learning. Finally, note that the words that had no gains were identified by more children at pretest than the words with gains. This suggests that, in some cases, lack of gains may have occurred because most of the children knew the word before the intervention began. For example, *solve* was identified by five of the children at pretest, and four of those same five (plus one additional child) also identified it at posttest. With a high level of performance at the pretest, it is not possible to demonstrate much gain on a 3AFC task.

Discussion

Rich vocabulary instruction embedded within a hands-on science curriculum was effective, to varying degrees, for five children but one child did not respond to the intervention and another could not complete the intervention. Below we explore potential reasons for these modest outcomes and propose steps towards finding solutions.

Some children with DLD do not respond to rich vocabulary instruction

The two children who did not demonstrate benefits from the intervention, Brianna and Tyrone, presented with triple risks: DLD, low SES, and poor executive function. Low SES can be viewed as a proxy for circumstances that threaten satisfaction of basic needs and stress the family, and both of these children and their families experienced multiple challenges during the time that camp was held. These two children were also the least engaged in the camp experience and the associated vocabulary intervention. Although this could have reflected a lack of interest, it is likely that their difficulties with language processing, working memory, and impulse control together with the stress they were experiencing meant that they could not engage consistently, even had interest been high.

It is useful to consider Brianna and Tyrone's lack of response in comparison to the best responder, Elvis. Like Brianna and Tyrone, Elvis lived in poverty and was in the midst of significant disruptions at home, but he had three advantages. First, he entered camp knowing more about science than any other participant. His good performance on the vocabulary pretest reflected that knowledge (but note that Brianna matched his score on the vocabulary pretest). Children demonstrate more robust learning of new words from semantic categories that are already well represented in their lexicons (Borovsky et al., 2016) so Elvis may have leveraged his superior knowledge of science and related words in a way that was unavailable to Brianna and Tyrone. Second, is the age difference; he was six and they were four. Sénéchal and Cornell (1993) found typical five-year-olds to benefit more from direct vocabulary instruction than four-year-olds and Dockrell et al. (2007) found age to be a positive predictor of word learning in response to video science-based instruction amongst typical children ages 4 to 7 years. Given his age advantage, he was likely more ready to learn. Third, he had intact executive function skills. Children with DLD, like Elvis, who bring good memory and impulse control to the task of word learning fare better than those who do not (Kapa & Erikson, 2020).

Note that the age advantage and the stronger executive function skills likely work together here. Executive function develops from infancy into young adulthood and, in particular, working memory systems controlled by the prefrontal cortex come on-line between four and eight years of age (Luciana, & Nelson, 1998). Because Brianna and Tyrone were four years old and Elvis was six, they would have had lower executive function skills even if they had good executive function skills for their age. They did not. Instead, they had poor executive function for their (younger) age whereas Elvis had strong executive function for his (older) age. Disentangling the effects of age and executive function is not possible here but that goal should be considered in future experimental work.

Previous studies linking a poor response to rich vocabulary instruction to the double risk of DLD + low SES have not measured executive function (Marulius & Neuman, 2010, 2013), but we know that children who are exposed to chronic poverty are at high risk for developing executive function deficits (Raver et al., 2013). It seems likely that the effect of poverty on vocabulary learning is mediated, in part, by its effect on the development of executive function. More research is needed to determine the extent to which low SES, executive function deficits, and their combination present risks to word learning and to establish the factors that may enable some learners, like Elvis, to remain resilient in the face of such risks.

Some children with DLD respond to rich vocabulary instruction but gains are modest

Five children benefitted from the intervention. After a therapeutically high dose for each of 20 words delivered across one-month, their accuracy improved between 15% (Luke) and 30% (Elvis), a gain in 3 to 6 words identified. To put these modest outcomes in perspective, the children were simultaneously learning other useful words, concepts, and processes during their science lessons, as would be true if they were enrolled in speech-language intervention while partaking in their school curriculum. Moreover, we were able to follow-up with Adam and Sanjay six months after the intervention ended and both had retained what they had learned and built upon it. This is especially encouraging given the severity of Adam's DLD.

The modest vocabulary outcomes that we obtained are not unusual. After an intervention involving 30 words with a dose of 36 exposures to each, Storkel et al. (2019) found that preschoolers with DLD had learned to define 9 or 10, on average, but remembered only 3 of them two weeks later. That intervention was not embedded in the curriculum. S. Gillam et al. (2014) taught first graders with or without DLD 40 words relevant to literacy (e.g., *author*) and the content of curricular books. They rated the children's definitions of those words pre- and post- a six-week intervention delivered in a classroom setting three times per week for 30 minutes per session. The children earned zero points for incorrect definitions, 1 point for an incomplete definition, and 2 points for a complete definition. The children with DLD gained 7 points, on average, from pretest to posttest. Given the scoring system, this would equate to learning seven new words to a minimal level, learning fewer than seven new words to a complete level, deepening knowledge of seven known words to a complete level, or some combination of these outcomes.

Rich vocabulary instruction is an evidence-based practice that incorporates long-attested principles of good teaching. In the current study, student clinicians exceeded a priori standards for the fidelity of rich instruction, but they did not use the full variety of recommended instructional strategies within the context of science lessons. We view the modest outcomes of this study, and those in Storkel et al. (2019) and S. Gillam et al. (2014), not as a reflection of the teaching approach *per se*, but a reflection of the challenge inherent in providing a full version of that approach to learners with varying needs within the context and constraints of academic instruction. Limited response to language learning opportunities is, in effect, a diagnostic sign of

DLD (Bishop et al., 2016). To experience gains, these children will need long-term, individually tailored interventions that are expertly delivered.

Individualizing rich vocabulary instruction

Individualizing the targets. The pressure to select words associated with lessons in the classroom is valid; the child will need these words to access the curriculum. However, words selected from the curriculum can be complex in form and meaning. Even typical learners may exhibit modest gains in science vocabulary learning opportunities, an outcome attributed, in part, to the conceptual complexity of the words (Dockrell et al., 2007). Some children with DLD will not respond well to a one-size-fits-all core vocabulary as we delivered here. These learners may need to target earlier acquired Tier 2 words or, in some cases, drop to Tier 1 words. SLPs include not only Tier 2 words but also Tier 1 words in vocabulary interventions for school children (Justice et al., 2014; Steele, 2019), and this may be appropriate for some learners.

Individualizing the focus. The phonological complexity of the word forms themselves, whether Tier 1 or 2, may be an obstacle for some learners with DLD. In the case studies reported here, longer words tended to be harder than shorter words. Likewise, Jackson et al. (2019) found that five-year-olds with DLD fast-mapped two-syllable words as well as their peers but they were significantly poorer at fast-mapping three- and four-syllable words. Longer words place greater demands on phonological working memory processes thought to support word learning (Gathercole & Baddeley, 1990; Kan & Windsor, 2010) and known to be limited in children with DLD (Archibald & Gathercole, 2007). Words comprised of later acquired phonemes may also be challenging. The articulatory demands of words containing late-8 sounds limit phonological working memory performance in adults (Moore et al., 2010), and, by extension, may exceed the memory span of children with DLD. That said, some children with DLD who have both language and speech sound problems will need to target late-8 sounds. Careful selection of word targets may leverage the interplay between phonological and lexical levels for more immediate gains in these children (Storkel, 2014). If words with complex or long word forms must be targeted in service of functionality or speech sound learning, a focus on awareness and production of phonological forms may be warranted.

Because a second intervention arm in our study involved phonological awareness training, we stripped strategies such as clapping syllabic sequences or emphasizing common onsets and rhymes from the rich vocabulary instruction that we provided. That decision is consistent with rich instruction approach as described in Beck et al. (2008, 2013), where depth of word meaning is emphasized over form. Nevertheless, people with DLD find it particularly difficult to learn word forms (McGregor et al., 2017), thus they may be especially in need of such instruction. Lowe et al. (2017) conducted a systematic review of vocabulary intervention studies for adolescents with DLD and concluded that interventions that combined semantic and phonological emphases were more effective than interventions that targeted either of these alone.

Individualizing the dose. Although all the words we targeted were modeled in therapeutically high doses, there was variability, and patterns in the words learned or not learned suggested that those given in higher doses were more often learned. Some children may need higher doses than others. In adults with DLD, doses required to add at least 13 of 15 new words to the expressive vocabulary ranged from 10 exposures per word for the best learner to more than 18 exposures per word for the weakest learners (McGregor et al., 2020). Providing the necessary dosage in a school environment, where direct contact with the SLP may be minutes per week, will require “all hands on deck,” not only the SLP but also the classroom teacher, resource teacher, and, if possible, the family (Ukrainetz & Fresquez, 2003). In fact, an advantage of selecting targets from the curriculum is that they will, naturally, be reinforced in the classroom. We take it as a good sign that the words more often learned in this study were those that had appeared more often in the FOSS scripts. The children may have found these more learnable not only because of the higher dosage but also their higher relevance to the science lessons.

The most typical model of service provision in U.S. public elementary schools is to pull a child from the classroom for individual or small group speech-language intervention (American Speech-Language-Hearing Association, 2020). Although rich vocabulary instruction can be provided in such a context, children typically receive one hour or less of pull-out intervention per week (American Speech-Language-Hearing Association, 2018), and this is not enough time to provide a therapeutic dose. Another approach is to embed the SLP in the classroom to provide comprehensive, integrated services alongside the teacher, a model known as interprofessional practice. Theoretically, this model could solve the low dosage problem and allow authentic contexts for the integration of rich vocabulary instruction. In fact, interprofessional collaboration with supplemental pull-out services yields better vocabulary outcomes for children with IEPs than pull-out only models or classroom-based models where the teacher and SLP work independently (plus supplemental pull-out) (Throneburg et al., 2000). However, only 14% of SLPs in U.S. elementary schools report using this model (Pfeiffer et al., 2019) and, among those who do, 54.2% engage in interprofessional treatment practice only occasionally, monthly or less (American Speech-Language-Hearing Association, 2016). Barriers to interprofessional collaboration include lack of training, lack of experience, and high caseload size (Throneburg et al., 2000). We conclude that the most frequently used models, pull-out or occasional-only interprofessional treatments, are likely to allow only a low-dose or partial version of rich vocabulary instruction that will not be enough for many learners with DLD.

Here we circle back to the nature of DLD. DLD is, by definition, hard to remediate. Peña et al. (2001) used brief word-instruction experiences as a dynamic assessment of children's language learning abilities and rated the efforts of both the child and the examiner during those experiences. Children with typical language not only had better outcomes than those with suspected DLD, but they also had different approaches to learning. The typical children were judged as "highly modifiable;" they were able to initiate and maintain attention to the task; they could plan and self-regulate and they stayed motivated if they made mistakes—all signs of intact executive function. The examiner exerted minimal effort in helping these children to achieve good outcomes. In contrast, learners with DLD did not stick with the task, and the gains that they did make required a high level of support and encouragement from the examiner. These dynamic assessment outcomes mirror our intervention outcomes. As a group, children with DLD do not learn language readily. Those who do respond well to language intervention are likely receiving extreme levels of support from their interventionist and are accessing the intervention experience by virtue of their stronger executive function skills.

Conclusions and future directions

Case series have inherent limitations in external validity. As we worked towards a rich description of these particular cases, we identified other limitations specific to this study. We wished that we had administered a comprehensive Adverse Childhood Experiences Survey (Center for Disease Control, 2020) to each parent to allow a more complete exploration of stressors in all families. Although we designed an intervention that ensured rich instruction, it did not include a focus on the phonological form of the target words, and this may have limited the outcomes. Also, a measure of vocabulary outcomes that allowed us to examine nuanced changes in depth of word knowledge would have been ideal. Our pilot data revealed that the children did not have the metalinguistic awareness to reflect on their knowledge of words, so we abandoned a 'levels of knowing' approach and used a 3AFC recognition task instead. We had hoped that the children's spontaneous productions would be another index of learning, but these were nearly nonexistent. Finally, because only three of the seven families were available for long-term follow-up, our retention outcomes are incomplete.

Case series also have inherent advantages in enabling understandings that are obscured by group means and parametric statistics. What this particular case series offered is an illustration of the challenge of translating vocabulary intervention research to practice. It also motivated several hypotheses:

- Individual children with DLD vary one to another in response to vocabulary intervention because of within-learner differences in interest, motivation, engagement, readiness to learn, executive function, and extant language abilities.
- Older children may respond better to rich vocabulary intervention than younger children.
- Children's responses to vocabulary intervention vary because of factors that exist outside the learner. These may include the skill of the clinician, the dosage, the conceptual and phonological complexity of the words to be learned, the relevance of those words to the context in which they are taught, and the fit between the child, the intervention approach, and the context in which it is delivered.
- Low SES is a proxy for numerous risks that directly or indirectly impede development and learning. These may include lack of basic needs (food, shelter, and health care), stressors on the family (illness, lack of support, exposure to crime), and reduced resources (time to devote to conversational interactions, money to pay for travel to intervention sessions).
- The modest success of rich vocabulary instruction for children with DLD is not a limitation of the approach itself but, rather, the difficulty of delivering such an intervention while tailoring the targets, approach, and dosage to the needs of individual children with DLD, children who vary one from the other in their linguistic strengths and weaknesses and in the presence and extent of accompanying deficits.

To facilitate translation of research on rich vocabulary instruction to practice, researchers should conduct randomized controlled trials in authentic contexts to determine not only that an intervention can be efficacious but that it can be effective outside of laboratory conditions. However, that alone will not be enough. A crucial next step will be to supplement randomized controlled trials—designs which are, by definition, one-size-fits-all—with methods that test the effects of tailoring rich vocabulary instruction to the needs of each learner. Carey and Stiles (2016) offer alternatives to randomized controlled trials for assessment of psychotherapies. Storkel, Komesidou et al. (2017) applied one of them, benchmarking, to determine which individual children with DLD needed a modification in their vocabulary intervention approach. This alternative design holds promise for testing tailored approaches.

We are not alone in calling for vocabulary interventions that are more precisely tailored to individual needs. When school SLPs were asked what would be required for them to deliver an ideal vocabulary intervention, they emphasized three changes. The first—broadening the scope of responsibility for vocabulary development via collaboration with professionals and parents—would go far towards satisfying necessary dosage levels. The second and third reflect the need for individualization: increasing the relevance of vocabulary intervention to each child; and finding more time to a) determine the strategies and approaches that work best for an individual child and b) to deliver individualized interventions by seeing children with common goals in smaller groups (Steele, 2020). Ensuring the *individualized* within a child's Individualized Education Program should be our common goal.

Acknowledgments

We are grateful for the support of NSF grant # 1748298 *Improving STEM Outcomes for Young Children with Language Learning Disabilities by Intervening at the Intersection of Language and Scientific Thought*, our participants and their parents, and the hard work of the following people.

- Fidelity tracking: Claire D'Antonio and Bonnie McNellis
- Materials preparation: Grace McInerney, Mitali Patel, Claire D'Antonio
- Science lesson design: Kelsey Newkirk
- Before and after care – Hannah Menella, Chris Elko, Gillian Isabelle, Beckey Carlson-Lee
- Transcription and data entry: Vanessa Coppola, Samantha Aquino
- The student clinicians
- Master Clinicians: Elizabeth Lucas, Meredith Bailey-Orr, and Renee Reilly
- Clinic support: Jacquiline Truluck and Christine Cook
- Access to physical space: The College School and the Lab Preschool at the University of Delaware
- General assistance and logistics: Kristina Strother Garcia
- Characterizing target words: Timothy Arbisi-Kelm
- Tracking of target vocabulary in teaching scripts: Kelsey McKernan
- Coding of circle-time interactions: Shivani Gajre
- Feedback on earlier draft: Erin Smolak, Ronald Pomper, Katherine Gordon, Hope Lancaster
- External Advisory Board for the project: Lizbeth Finestack, Kathy Long, Mary Lukas, and Holly Storkel.

References

Alt, M., & Plante, E. (2006). Factors that influence lexical and semantic fast mapping of young children with specific language impairment. Journal of Speech, Language, and Hearing Research, 49(5), 941–954. https://doi.org/10.1044/1092-4388 (2006/068)

American Speech-Language-Hearing Association. (n.d.). Childhood hearing screening. https://www.asha.org/practice-portal/ professional-issues/childhood-hearing-screening/#collapse_1

American Speech-Language-Hearing Association. (2016). Interprofessional practice survey results. https://www.asha.org/uploadedFiles/ 2016-Interprofessional-PracticeSurvey-Results.pdf

American Speech-Language-Hearing Association. (2018). 2018 Schools Survey report: SLP caseload and workload characteristics. https://www.asha.org/uploadedFiles/Schools-2018-SLP-Caseload-and-Workload-Characteristics.pdf

American Speech-Language-Hearing Association. (2020). 2020 Schools Survey report: SLP caseload and workload characteristics. https://www.asha.org/uploadedFiles/2020-Schools-Survey-SLP-Caseload.pdf

Archibald, L. M., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders, 41*(6), 675-693. https://doi.org/10.1080/13682820500442602

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal, 107*(3), 251-271. https://www.journals.uchicago.edu/doi/abs/10.1086/511706

Beck, I. L., McKeown, M. G., & Kucan, L. (2008). *Creating robust vocabulary: Frequently asked questions and extended examples*. Guilford Press.

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. Guilford Press.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology, 93*(3), 498. https://psycnet.apa.org/doiLanding?doi=10.1037%2F0022-0663.93.3.498

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & The Catalise Consortium. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLOS one, 11*(7), e0158753. https://doi.org/10.1371/journal.pone.0158753

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Lexical leverage: Category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science, 19*(6), 918-932. https://doi.org/10.1111/desc.12343

Brigham, F. J., Scruggs, T. E., & Mastropieri, M. A. (2011). Science education and students with learning disabilities. *Learning Disabilities Research & Practice, 26*(4), 223-232. https://doi.org/10.1111/j.1540-5826.2011.00343.x

Carey, T. A., & Stiles, W. B. (2016). Some problems with randomized controlled trials and some viable alternatives. *Clinical Psychology & Psychotherapy, 23*(1), 87-95. https://doi.org/10.1002/cpp.1942

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., Lively, T.

J., & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188-215. https://doi.org/10.1598/RRQ.39.2.3

Center for Disease Control (2020). Adverse childhood experiences (ACES). https://www.cdc.gov/violenceprevention/acestudy/ . Downloaded 8/23/20.

Chung, S. F. (2012). Research based vocabulary instruction for English Language Learners. *The Reading Matrix, 12*(2), 105-120. http://vocabforell.pbworks.com/w/file/fetch/92126205/Chung.pdf

Common Core Standards Initiative. Key shifts in language arts. http://www.corestandards.org/other-resources/key-shifts-in-english-language-arts/. Downloaded 8/10/2020.

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli Jr, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: *Teaching for breadth versus depth. The Elementary School Journal, 110*(1), 1-18. https://www.journals.uchicago.edu/doi/abs/10.1086/598840

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934. https://doi.org/10.1037/0012-1649.33.6.934

Deevy, P., Weil, L. W., Leonard, L. B., & Goffman, L. (2010). Extending use of the NRT to preschool-age children with and without specific language impairment. *Language, Speech, and Hearing Services in Schools, 41*(3), 277-288. https://doi.org/10.1044/0161-1461(2009/08-0096)

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959-964. https://doi.org/10.1126/science.1204529

Dockrell, J. E., Braisby, N., & Best, R. M. (2007). Children's acquisition of science terms: Simple exposure is insufficient. *Learning and Instruction, 17*(6), 577-594. https://doi.org/10.1016/j.learninstruc.2007.09.005

Dockrell, J. E., Lindsay, G., Connelly, V., & Mackie, C. (2007). Constraints in the production of written text in children with specific language impairments. *Exceptional Children, 73*(2), 147-164. https://doi.org/10.1177/001440290707300202

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136-1146. https://doi.org/10.1044/jslhr.4105.1136

Dunn, L. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test-IV*. Pearson.

Ebert, K. D., & Kohnert, K. (2011). Sustained attention in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 54*(5):1372–1384. https://doi.org/10.1044/1092-4388(2011/10-0231)

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*(3), 393-447. https://doi.org/10.3102/00346543071003393

Elliot, C. D. (2007). *Differential Ability Scales-II (DAS-II)*. San Antonio, TX: Pearson Clinical.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection?. *Journal of Memory and Language, 29*(3), 336-360. https://doi.org/10.1016/0749-596X(90)90004-J

Gillam, R. B., & Pearson, N. (2004). *Test of Narrative Language*. Austin, TX: Pro-Ed.

Gillam, S.L. , Olszewski, A., Fargo, J., & Gillam, R. B. (2014). Classroom-based narrative and vocabulary instruction: Results of an early-stage, nonrandomized comparison study. *Language, Speech, and Hearing Services in Schools, 45*(3), 204-219. https://doi.org/10.1044/2014_LSHSS-13-0008

Gioia, G. A., Espy, K. A, & Isquith, P. K. (2003). *Behavior Rating Inventory of Executive Function- Preschool Version*. Lutz FL: PARINC.

Goldstein, H., Ziolkowski, R. A., Bojczyk, K. E., Marty, A., Schneider, N., Harpring, J., & Haring, C. D. (2017). Academic vocabulary learning in first through third grade in low-income schools: Effects of automated supplemental instruction. *Journal of Speech, Language, and Hearing Research, 60*(11), 3237-3258. https://doi.org/10.1044/2017_JSLHR-L-17-0100

Golinkoff, R. M., De Villiers, J. G., Hirsh-Pasek, K., Iglesias, A., Wilson, M. S., Morini, G., & Brezack, N. (2017). *User's manual for the Quick Interactive Language Screener (QUILS): A measure of vocabulary, syntax, and language acquisition skills in young children*. Paul H. Brookes Publishing Company.

Gonzalez, J. E., Pollard-Durodola, S., Simmons, D. C., Taylor, A. B., Davis, M. J., Kim, M., & Simmons, L. (2010). Developing low-income preschoolers' social studies and science vocabulary knowledge through content-focused shared book reading. *Journal of Research on Educational Effectiveness, 4*(1), 25-52. https://doi.org/10.1080/19345747.2010.487927

Gordon, K. R. (2020). The advantages of retrieval-based and spaced practice: Implications for word learning in clinical and educational contexts. *Language, Speech, and Hearing Services in Schools, 51*(4), 1-11. https://doi.org/10.1044/2020_LSHSS-19i-00001

Gray, S. (2003). Word-learning by preschoolers with specific language impairment: What predicts success?. *Journal of Speech, Language, and Hearing Research, 46(1)*:56–67. https://doi.org/10.1044/1092-4388(2003/005)

Gray, S. (2004). Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research, 47*(5):1117–1132. https://doi.org/10.1044/1092-4388(2004/083)

Haebig, E., Leonard, L. B., Deevy, P., Karpicke, J., Christ, S. L., Usler, E., Kueser, J. B., Souto, S., Krok, W., & Weber, C. (2019). Retrieval-based word learning in young typically developing children and children with development language disorder II: A comparison of retrieval schedules. *Journal of Speech, Language, and Hearing Research, 62*(4), 944-964. https://doi.org/10.1044/2018_JSLHR-L-18-0071

Jackson, E., Leitao, S., Claessen, M., & Boyes, M. (2019). Fast mapping short and long words: Examining the influence of phonological short-term memory and receptive vocabulary in children with developmental language disorder. *Journal of Communication Disorders, 79*, 11-23. https://doi.org/10.1016/j.jcomdis.2019.02.001

Johnson, G., Gersten, R., & Carnine, D. (1987). Effects of instructional design variables on vocabulary acquisition of LD students: A study of computer-assisted instruction. *Journal of Learning Disabilities, 20*(4), 206-213. https://doi.org/10.1177/002221948702000402

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks. *Language, Speech, and Hearing Services in Schools, 36*(1), 17–32. https://doi.org/10.1044/0161-1461(2005/003)

Justice, L. M., Schmitt, M. B., Murphy, K. A., Pratt, A., & Biancone, T. (2014). The 'robustness' of vocabulary intervention in the public schools: targets and techniques employed in speech–language therapy. *International Journal of Language & Communication Disorders, 49*(3), 288-303. https://doi.org/10.1111/1460-6984.12072

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 53*(3), 739–756. https://doi.org/10.1044/1092-4388(2009/08-0248)

Kapa, L. L., & Erikson, J. A. (2020). The relationship between word learning and executive function in preschoolers with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research, 63*(7), 1-15. https://doi.org/10.1044/2020_JSLHR-19-00342

Karpicke, J. D., & Bauernschmidt, A. (2011). Spaced retrieval: absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1250. https://doi.org/10.1037/a0023436

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966-968. https://doi.org/10.1126/science.1152408

Kong, N. Y. (2013). *The Effectiveness of a Supplemental Pre-Kindergarten Vocabulary Intervention* (Doctoral dissertation, University of Kansas). http://hdl.handle.net/1808/15116

Lawrence Hall of Science. (2015). *Full Option Science System* [Series]. Delta Education.

Loftus, S. M., & Coyne, M. D. (2013). Vocabulary instruction within a multi-tier approach. *Reading & Writing Quarterly, 29*(1), 4-19. https://doi.org/10.1080/10573569.2013.741942

Loftus, S. M., Coyne, M. D., McCoach, D. B., Zipoli, R., & Pullen, P. C. (2010). Effects of a supplemental vocabulary intervention on the word knowledge of kindergarten students at risk for language and literacy difficulties. *Learning Disabilities Research & Practice, 25*(3), 124-136. https://doi.org/10.1111/j.1540-5826.2010.00310.x

Lowe, H., Henry, L., Müller, L. M., & Joffe, V. L. (2018). Vocabulary intervention for adolescents with language disorder: a systematic review. *International Journal of Language & Communication Disorders, 53*(2), 199-217. https://doi.org/10.1111/1460-6984.12355

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four-to eight-year-old children. *Neuropsychologia, 36*(3), 273-293. https://doi.org/10.1016/S0028-3932(97)00109-7

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research, 80*(3), 300-335. https://doi.org/10.3102/0034654310377087

Marulis, L. M., & Neuman, S. B. (2013). How vocabulary interventions affect young children at risk: A meta-analytic review. *Journal of Research on Educational Effectiveness, 6*(3), 223-262. https://doi.org/10.1080/19345747.2012.755591

Marzano, R. J., & Rogers, K. (2012). *Vocabulary for the new science standards*. Solution Tree Press.

Maynard, K. L., Pullen, P. C., & Coyne, M. D. (2010). Teaching vocabulary to first-grade students through repeated shared storybook reading: A comparison of rich and basic instruction to incidental exposure. *Literacy Research and Instruction, 49*(3), 209-242. https://doi.org/10.1080/19388070902943245

McGregor, K. K. (2020, August 24). The Challenge of Rich Vocabulary Instruction for Children

with Developmental Language Disorder. Retrieved from osf.io/ytzjp
McGregor, K. K., Arbisi-Kelm, T., & Eden, N. (2017). The encoding of word forms into memory may be challenging for college students with developmental language impairment. *International Journal of Speech-Language Pathology, 19*(1), 43-57. https://doi.org/10.3109/17549507.2016.1159337
McGregor, K. K., Arbisi-Kelm, T., Eden, N., & Oleson, J. (2020). The word learning profile of adults with developmental language disorder. *Autism & Developmental Language Impairments, 5*, 2396941519899311. https://doi.org/10.1177/2396941519899311
McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A. J., & Lloyd, M. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *Journal of Autism and Developmental Disorders, 42*(1), 35-47. https://doi.org/10.1007/s10803-011-1210-4
McGregor, K. K., & Duff, D. (2015). Promoting diverse and deep vocabulary development. In: T. Ukrainetz, Editor, *School-age language intervention: Evidence based practices*, (pp. 247-278). Austin, TX: Pro-Ed.
McGregor, K. K., Gordon, K., Eden, N., Arbisi-Kelm, T., & Oleson, J. (2017). Encoding deficits impede word learning and memory in adults with developmental language disorders. *Journal of Speech, Language, and Hearing Research, 60*(10), 2891-2905. https://doi.org/10.1044/2017_JSLHR-L-17-0031
McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. Journal of Speech, Language, and Hearing Research, 45(5), 998–1014. https://doi.org/10.1044/1092 4388(2002/081)
McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language & Communication Disorders, 48*(3), 307-319. https://doi.org/10.1111/1460-6984.12008
Moore, M. W., Tompkins, C. A., & Dollaghan, C. A. (2010). Manipulating articulatory demands in non-word repetition: a 'late-8'non-word repetition task. *Clinical Linguistics & Phonetics, 24*(12), 997-1008. https://doi.org/10.3109/02699206.2010.510917
National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.
Nowicki, B. L., Sullivan-Watts, B., Shim, M. K., Young, B., & Pockalny, R. (2013). Factors influencing science content accuracy in elementary inquiry science lessons. *Research in Science Education, 43*(3), 1135-1154. https://doi.org/10.1007/s11165-012-9303-4
Pauls, L. J., & Archibald, L. M. (2016). Executive functions in children with specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 59*(5), 1074-1086. https://doi.org/10.1044/2016_JSLHR-L-15-0174
Pearson, Inc. (2018). PPVT-5 and EVT-3 By the Numbers. https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/ppvt-5/ppvt-5-and-evt-3-technical-summary.pdf. Downloaded 8/3/2020.
Peña, E. D., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of

children's word learning ability. *American Journal of Speech Language Pathology, 10*(2), 138–154. https://doi.org/10.1044/1058-0360(2001/014)

Perona, K., Plante, E., & Adam, R. (2005). Diagnostic Accuracy of the Structured Photographic Expressive Language Test: Third Edition (SPELT®-3). *Language, Speech, and Hearing Services in Schools, 36*(2), 103-115. https://doi.org/10.1044/0161-1461(2005/010)

Raver, C. C., Blair, C., & Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: New perspectives on models of differential susceptibility. *Developmental Psychology, 49*(2), 292. https://doi.org/10.1037/a0028343

Rutter, M., Bailey, A., & Lord, C. (2003). *The social communication questionnaire*. Los Angeles: Western Psychological Services.

Schickedanz, J., & Dickinson, D. (2005). *Opening the World of Learning*. Iowa City, IA: Pearson Publishing.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals Fourth Edition (CELF-4)*. Bloomington, MN: Pearson.

Sénéchal, M. and Cornell, E. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly, 28*(4): 360–374. https://doi.org/10.2307/747933

Seymour, H. N., Roeper, T. W., & de Villiers, J. (2003a). *Diagnostic evaluation of language variation: Criterion referenced*. San Antonio, TX: Psychological Corporation.

Seymour, H. N., Roeper, T. W., & De Villiers, J. (2003b). *Diagnostic evaluation of language variation: Screening test*. San Antonio, TX: Psychological Corporation.

Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders I: A clinical profile. *Journal of Speech and Hearing Research, 37*(5), 1100–1126. https://doi.org/10.1044/jshr.3705.1100

Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology, 101*(2), 305. https://doi.org/10.1037/a0014217

Smolak, E., McGregor, K. K., Arbisi-Kelm, T., & Eden, N. (2020). Sustained attention in developmental language disorder and its relation to working memory and language. *Journal of Speech, Language, and Hearing Research*.

Steele, S. C. (2020). Vocabulary intervention: A national survey of school-based speech language pathologists. *Communication Disorders Quarterly, 41*(3), 151-161. https://doi.org/10.1177/1525740119827008

Steele, S. C., & Mills, M. T. (2011). Vocabulary intervention for school-age children with language impairment: A review of evidence and good practice. *Child Language Teaching and Therapy, 27*(3), 354-370. https://doi.org/10.1177/0265659011412247

Storkel, H. L. (2018). Implementing evidence-based practice: Selecting treatment words to boost phonological learning. Language, Speech, and Hearing Services in Schools, 49(3), 482–496. https://doi.org/10.1044/2017_LSHSS-17-0080

Storkel, H. L., Komesidou, R., Fleming, K. K., & Romine, R. S. (2017). Interactive book reading to accelerate word learning by kindergarten children with specific language impairment: Identifying adequate progress and successful learning patterns. *Language, Speech, and Hearing Services in Schools, 48*(2), 108-124. https://doi.org/10.1044/2017_LSHSS-16-0058

Storkel, H. L., Komesidou, R., Pezold, M. J., Pitt, A. R., Fleming, K. K., & Romine, R. S.

(2019). The impact of dose and dose frequency on word learning by kindergarten children with developmental language disorder during interactive book reading. *Language, Speech, and Hearing Services in Schools, 50*(4), 518-539. https://doi.org/10.1044/2019_LSHSS-VOIA-18-0131

Storkel, H. L., Voelmle, K., Fierro, V., Flake, K., Fleming, K. K., & Romine, R. S. (2017). Interactive book reading to accelerate word learning by kindergarten children with specific language impairment: Identifying an adequate intensity and variation in treatment response. *Language, Speech, and Hearing Services in Schools, 48*(1), 16–30. https://doi.org/10.1044/2016_LSHSS-16-0014

Stothard, S. E., Snowling, M. J., Bishop, D. V., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41*(2), 407-418. https://doi.org/10.1044/jslhr.4102.407

Therrien, W. J., Taylor, J. C., Hosp, J. L., Kaldenberg, E. R., & Gorsh, J. (2011). Science instruction for students with learning disabilities: A meta-analysis. *Learning Disabilities Research & Practice, 26*(4), 188-203. https://doi.org/10.1111/j.1540-5826.2011.00340.x

Throneburg, R. N., Calvert, L. K., Sturm, J. J., Paramboukas, A. A., & Paul, P. J. (2000). A comparison of service delivery models: Effects on curricular vocabulary skills in the school setting. *American Journal of Speech-Language Pathology, 9*(1), 10-20. https://doi.org/10.1044/1058-0360.0901.10

Ukrainetz, T. A., & Fresquez, E. F. (2003). What isn't language?. *Language, Speech, and Hearing Services in Schools, 34*(4): 284-298. https://doi.org/10.1044/0161-1461(2003/024)

Weiland, C., Barata, M. C., & Yoshikawa, H. (2014). The co-occurring development of executive function skills and receptive vocabulary in preschool-aged children: A look at the direction of the developmental pathways. *Infant and Child Development, 23*(1), 4-21. https://doi.org/10.1002/icd.1829

Williams, K. T. (2007). *Expressive vocabulary test second edition*. San Antonio, TX: Pearson.

Windfuhr, K. L., Faragher, B., & Conti-Ramsden, G. (2002). Lexical learning skills in young children with specific language impairment (SLI). *International Journal of Language & Communication Disorders, 37*(4), 415-432. 432. https://doi.org/10.1080/1368282021000007758

Young, B. J., & Lee, S. K. (2005). The effects of a kit-based science curriculum and intensive science professional development on elementary student science achievement. *Journal of Science Education and Technology, 14*(5-6), 471-481. https://doi.org/10.1007/s10956-005-0222-2

Zipoli, R. P., Coyne, M. D., & McCoach, D. B. (2011). Enhancing vocabulary intervention for kindergarten students: Strategic integration of semantically related and embedded word review. *Remedial and Special Education, 32*(2), 131-143. https://doi.org/10.1177/0741932510361262