ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

Evaluating the Efficacy of a Narrative Language Intervention for Bilingual Students

Philip Capin,[a] Sharon Vaughn,[a] Sandra Laing Gillam,[b] Anna-Maria Fall,[a] Gregory Roberts,[a] Megan Israelsen-Augenstein,[c] Sarai Holbrook,[d] Rebekah Wada,[e] Jordan Dille,[f] Colby Hall,[g] and Ronald B. Gillam[b]

[a]Department of Special Education, The University of Texas at Austin [b]Department of Communicative Disorders and Deaf Education, Utah State University, Logan [c]Department of Human Performance, School of Medicine, West Virginia University, Morgantown [d]Department of Communication Sciences and Disorders, University of Wisconsin–Stevens Point [e]School of Health Sciences, Francis Marion University, Florence, SC [f]Teacher Education, University of Nebraska at Kearney [g]Department of Curriculum, Instruction, and Special Education, University of Virginia, Charlottesville

ARTICLE INFO

Article History:
Received June 21, 2021
Revision received February 8, 2023
Accepted August 8, 2023

Editor-in-Chief: Erinn H. Finke
Editor: Anny Castilla-Earls

https://doi.org/10.1044/2023_AJSLP-21-00185

ABSTRACT

Purpose: This study examines the narrative language and reading outcomes of monolingual and bilingual students who received instruction with the Supporting Knowledge in Language and Literacy (SKILL) program, a narrative language intervention.
Method: The main effects of the SKILL program were evaluated in a randomized controlled trial in which students (N = 355) who were at risk for English language and literacy difficulties were randomized to the SKILL intervention or a business-as-usual instruction. This article reports secondary analyses examining the efficacy of SKILL for bilingual (n = 148) and monolingual (n = 207) students who completed measures of oral and written narrative language and reading comprehension in English.
Results: Moderation results showed that the effects of SKILL did not differ for monolinguals and bilinguals across most narrative language measures and did not vary for monolinguals or bilinguals based on their pre-intervention language performance.
Conclusion: These findings that suggest a language-based approach to improving narrative production and comprehension yielded similar results for monolinguals and bilinguals and that neither monolinguals nor bilinguals in this study needed to meet a certain threshold of English language proficiency to benefit from the intervention.

There are many benefits to being bilingual. Children who are bilingual are better able "to develop and sustain strong ties with their immediate and extended families" who do not speak English and "establish a strong cultural identity" (Espinosa, 2006, p. 2). In the long term, there may also be economic benefits to bilingualism as employers increasingly favor employees who are multilingual, contribute cultural expertise, and can work collaboratively with colleagues across ethnic and cultural backgrounds. This may be particularly true for this generation of children who will likely enter adulthood in a country that is a "majority–minority" society (e.g., Gándara, 2015). Bilingual students also contribute a myriad cultural and linguistic knowledge that teachers and students can access. Unfortunately, bilingual children in the United States disproportionately experience elevated levels of poverty and other sociocultural factors that make them vulnerable to academic difficulties. Spanish-speaking bilingual students, the largest subgroup of bilinguals in the United States, are more likely than monolingual peers to have family incomes below or near poverty levels (Fry & Gonzales, 2008; Hernandez et al., 2008) and parents with relatively low levels of education and literacy (Capps

Correspondence to Philip Capin: pcapin@utexas.edu. Jordan Dille is now at the University of Nebraska at Kearney. ***Disclosure:*** *Ronald B. Gillam receives royalties from the sale of the Test of Narrative Language, which was administered to the participants. Ronald B. Gillam and Sandra Laing Gillam receive royalties from the sale of the Supporting Knowledge in Language and Literacy intervention. All other authors have declared that no other competing financial or nonfinancial interests existed at the time of publication.*

et al., 2005; Hernandez et al., 2008). Due to policies outside of schools that lead to school funding inequities that negatively impact Latinos/Latinas and other ethnic minorities (e.g., Rothstein, 2017), they are also more likely to be enrolled in underresourced schools (Cosentino de Cohen et al., 2005). As a result, many bilingual students in the United States have fewer opportunities to access texts and academic experiences that contribute to successful academic English language and reading proficiency. These socio-contextual challenges are reflected in the high number (67%) of emergent bilingual (EB) learners, who are in the process of developing proficiency in English as well as another language, in the United States who score below the basic reading level in fourth grade (National Assessment of Educational Progress [NAEP], 2022).

Identifying evidence-based practices for addressing the English language literacy needs of EBs is of high importance to speech-language pathologists (SLPs) and reading specialists in the United States due to the increasing prevalence of bilinguals and the vulnerabilities these students experience (Cho et al., 2019). S. L. Gillam et al. (2023) recently reported findings from a large-scale (N = 356) randomized controlled trial (RCT) showing the positive effects of a narrative intervention program, namely, Supporting Knowledge in Language and Literacy (SKILL; S. L. Gillam et al., 2018), in improving outcomes for Grades 1–4 children with English language and literacy difficulties. The results from the analyses of the main effects (S. L. Gillam et al., 2023) revealed that the SKILL intervention was implemented with high fidelity and that students randomized to receive the SKILL intervention significantly outperformed students in a business-as-usual (BAU) control group on a variety of standardized and experimental measures of oral and written narrative language (g = .20–.61). Generalization to reading comprehension was not statistically significant for the entire sample when measured using a generalized measure of reading comprehension (Gates–MacGinitie Reading Test [GMRT] Reading Comprehension [RC] subtest; MacGinitie et al., 2001). However, results revealed that the experimental effect of SKILL on reading comprehension varied for students in Grades 1 and 2 relative to students in Grades 3 and 4, with older students randomized to SKILL showing greater performance (g = .26) than students in a BAU instruction.

These findings underscore the positive effects of SKILL in improving oral and written narrative language outcomes as well as the potential for these improvements to lead to gains in reading comprehension for upper elementary students. Although the SKILL program was originally developed with consideration for the needs of EBs, S. L. Gillam et al. (2023) did not address questions about the differential efficacy of SKILL for bilingual students in their sample (n = 148). This study extends the prior study on main effects by examining the extent to which the effects of the narrative intervention program vary for bilingual and monolingual students who are at risk for English language and reading comprehension difficulties.

The findings from this study will be particularly salient for SLPs because identifying empirically validated Tier 2 narrative language interventions has the potential to prevent language impairments. This is particularly critical in light of the shortage of SLPs in U.S. schools (American Speech-Language-Hearing Association [ASHA], 2018), which impacts the number of students who receive speech-language services and the quality of these services (Farquharson et al., 2021). It is also of high importance to SLPs because identification of communication disorders among bilingual students is particularly challenging (e.g., McLeod et al., 2017). Identifying evidence-based Tier 2 practices that can be used widely in schools may eventually help SLPs use response-to-instruction information to inform decisions about the identification of communication disorders.

Importance of Oral Language and Narrative Language Proficiency in Bilingual Students

Oral language is the linchpin to learning broadly as well as specifically for reading and literacy-related tasks for students, regardless of their primary language status. Oral language development is linked to literacy because it involves acquiring, practicing, and integrating phonology (sound units), semantics (vocabulary/academic language), morphology (grammatical morphemes), syntax (grammar), and pragmatics (social discourse skills; Catts et al., 2012). Oral narration, the ability to understand and tell stories, is related to reading comprehension because early reading instruction occurs in the context of narrative genre (Cook & O'Brien, 2014).

Narratives consist of story grammar elements (initiating events, internal responses, plans, attempts, consequences, and reactions) that represent the structure of episodes (commonly referred to as the macrostructure) and the words and sentences that form the story (i.e., microstructure; Greenhalgh & Strong, 2001; Mandler & Johnson, 1977; Stein & Glenn, 1979). Across languages and cultural groups, there are similarities in the story grammar elements that appear in children's stories (Berman & Slobin, 2013; Heilmann et al., 2016; McCabe & Bliss, 2005; Squires et al., 2014). However, the way story elements are sequenced—as well as the specific vocabulary and sentence structures in stories—often varies as a function of linguistic and sociocultural differences (Champion et al., 2003; S. L. Gillam et al., 2012; Iluz-Cohen & Walters, 2012; Price et al., 2006; Silliman et al., 2002; Simon-Cereijido & Gutiérrez-Clellen, 2009).

Narrative comprehension and production expectations are well represented as objectives for language and literacy instruction for elementary-age children (e.g., National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Expert recommendations for supporting reading comprehension in the elementary grades also identify teaching narrative language skills as a key element to supporting reading comprehension (Foorman et al., 2016). Narrative interventions are also frequently identified as one of the most impactful ways for school-based SLPs to improve language (e.g., Spencer & Petersen, 2020). Even though the development of narrative language ability is foundational to the English literacy development of bilingual students (August & Shanahan, 2006; Oller & Pearson, 2002), few studies have explored how it develops in this population. Spanish–English dual language learners (DLLs) have been shown to make similar gains in Spanish and English in their use of story elements as their language abilities improve. Pearson (2002) conducted a cross-sectional study with 240 typically developing monolingual and bilingual children in second and fifth grades. Students were asked to tell the story from the *Frog, Where Are You?* (Mayer, 2003) wordless picture book in Spanish and English (for the bilingual children). Overall narrative quality and level of language ability were calculated for all of the students. There was a significant interaction between grade and bilingual status (i.e., monolingual vs. bilingual language status), with the bilingual children scoring significantly lower than their monolingual peers in second grade. However, at fifth grade, narrative scores were not significantly different.

English narrative ability is an area of vulnerability for bilingual students (August et al., 2005; Oller & Pearson, 2002; Uccelli & Páez, 2007). This may be because their English language abilities are still developing and are not stable enough to support the cognitive and linguistic demands required to comprehend and compose narrative discourse. Prior research shows that stories told by bilingual students with developing English language and literacy tend to contain more ungrammatical utterances (Cleave et al., 2010), are shorter and less complex (McCabe & Bliss, 2005; Squires et al., 2014), and contain less complex vocabulary (Iluz-Cohen & Walters, 2012) compared to stories told by bilingual students with more developed language and literacy proficiency. Conversely, improvements in English narrative skills are well documented as bilingual students gain familiarity and experience with the English language (e.g., Lucero, 2018; J. F. Miller et al., 2006; Uccelli & Páez, 2007).

Oral Narrative Instruction for Monolinguals and Bilinguals

Recent systematic reviews of narrative language interventions (Favot et al., 2021; Nicolopoulou & Trapp, 2018; Pesco & Gagné, 2017; Petersen, 2011; Pico et al., 2021; Rogde et al., 2019) reveal that interventions that incorporate macrostructure instruction (i.e., instruction that teaches students narrative text structure through story grammar elements) demonstrate a positive impact on narrative language outcomes for preschool and elementary-age children with developing English language skills, including bilingual students. A recently published study by Petersen et al. (2022) was conducted to examine narrative outcomes of 686 kindergarten students who received whole-classroom narrative instruction twice weekly for 15–20 min. A secondary, quasi-experimental aspect of the study involved matching students (n = 49) from the original treatment and control groups who were identified as "at risk" to receive additional treatment. In this small cluster-randomized study, the authors reported statistically significant oral and written narrative and expository outcomes for the children who received additional treatment. The study did not include information about bilingual status or whether their outcomes varied as a function of bilingual status.

Few studies have examined the effects of oral narrative language instruction on language and literacy outcomes in bilingual school-age children directly. Spencer et al. (2020) conducted a cluster-randomized study to determine the effects of a dual language (Spanish–English) narrative intervention that incorporated macrostructure instruction for preschool children in Head Start schools with a home language of Spanish (N = 81). The authors reported statistically significant differences on a variety of investigator-designed measures of narrative production and vocabulary in English and Spanish, story comprehension in English, and sentence structure in English and Spanish that were closely aligned with the intervention. The study supports the feasibility and promise of a dual language narrative intervention for bilingual preschoolers.

A previous study by Spencer et al. (2015) compared outcomes of a narrative language intervention in a quasi-experimental study conducted in English with monolingual and bilingual students. The authors examined differential response to a narrative instruction among preschool monolingual (English-speaking) and bilingual children (N = 71). Preschoolers participated in story-retelling activities using illustrations and icons. Statistically significant differences were reported between students in treatment and control groups on researcher-designed measures of story retell and comprehension that were closely aligned with the intervention. Children who were DLLs did not demonstrate a different pattern of response to intervention compared to preschoolers who were monolingual English speakers. This is a promising result worthy of further exploration in large-scale studies and with older school-age students.

Unfortunately, the investigations conducted with bilingual school-age children have often been limited by small sample sizes and the use of nonrandomized designs (R. D. Miller et al., 2018; Peña et al., 2014). We were also unable to find any study that directly compared the outcomes of an oral narrative intervention conducted in English on monolingual and bilingual school-age children who were at risk for English language and literacy difficulties. Perhaps the largest study examining the effects of narrative instruction was conducted by the Language and Reading Research Consortium (LARRC; LARRC et al., 2019). The authors conducted an RCT (N = 938) with school-age children in Grades 1–3, including monolingual and bilingual students. They sought to understand the effects of *Let's Know!*—a language-focused supplemental intervention that targets children's language skills, including narrative and expository text structure knowledge, inference making, comprehension monitoring, and vocabulary. The LARRC reported statistically significant main effects on proximal measures of vocabulary and comprehension monitoring. There was also a small but statistically significant effect on narrative comprehension for students in Grade 3, but not for students in Grades 1 and 2. It is important to note that greater than 90% of the students in the study spoke a home language of English, but the authors did not examine whether there were differences in response to treatment based on bilingual status.

SKILL Program

To address the need to improve narrative language and literacy outcomes among elementary students at risk for English language and literacy difficulties, S. L. Gillam et al. (2018) developed the SKILL intervention as a supplemental language intervention that could be implemented within multitiered systems of support for children with or at risk for English language and literacy difficulties. The SKILL intervention program was developed with the primary intent of improving narrative language abilities for students who require additional oral language supports by experts in speech-language pathology in collaboration with school-based SLPs and special and general education teachers in a principled, study-by-study fashion with each pilot study informing the next (for a review of studies, see S. L. Gillam et al., 2018). This multiyear development work culminated in a theoretically grounded and practitioner-friendly manualized program that targets aspects of narrative macrostructure and microstructure.

The SKILL program addresses key elements of narratives through instruction targeting macrostructural aspects (characters, settings, initiating events, plans, attempts, internal responses, complications, consequences, and story endings; i.e., story grammar elements based on the episode structure in Stein & Glenn, 1979), with an explicit focus on the causal relationships between macrostructural and microstructural aspects of narratives (i.e., coordinated and subordinated conjunctions, adverbial and relative clauses, elaborated noun phrases, metacognitive and metalinguistic verbs) in the context of authentic narrative discourse. S. L. Gillam et al. (2018) designed SKILL to help students construct a mental model of narrative discourse at a global level (macrostructure) and to align with Kintsch's (2018) construction-integration model of comprehension. The SKILL program helps students construct microstructures, a textbase of information that connects to the macrostructure. An important focus of the program is its attention to the nature of causal relationships during macrostructure instruction. Rather than introduce and practice each story element in isolation, lessons promote the use of causal adverbs (e.g., "because" and "so") to tie the elements together causally.

There are a few reasons why SKILL may be particularly promising for bilingual students with developing English language and literacy skills. First, SKILL addresses a broad constellation of academic language targets (e.g., morphology, syntax, semantics/vocabulary) that underlie text comprehension (Uccelli et al., 2015). It provides children an opportunity to learn to integrate linguistic knowledge across domains (semantic, syntactic, morphologic, pragmatic) during spoken discourse, which is the modality in which children initially develop these skills. Second, S. L. Gillam et al. (2018) developed the SKILL program to meet the needs of students who were at risk for English language and literacy problems, with a specific focus on bilingual students by aligning the program with evidence-based practices for teaching academic content and literacy to bilinguals in the elementary grades, including (a) teaching vocabulary using direct instruction and in the context of stories; (b) defining new words and concepts by linking their meanings to familiar concepts in interactive discussions; (c) providing visual and gestural cues; (d) providing small-group instruction to distribute working memory load; and (e) providing opportunities for repetitive, distributive practice over multiple sessions (Armon-Lotem et al., 2021; Baker et al., 2014; Bedore et al., 2020; Kong & Hurless, 2023; Méndez et al., 2015; Restrepo et al., 2013; Spencer et al., 2020; Thordardottir, 2010; Ukrainetz, 2006).

Consistent findings from a series of small-scale pilot studies with low-achieving students at risk for disabilities (S. L. Gillam et al., 2014), EBs at risk for disabilities (S. L. Gillam et al., 2018; Jensen, 2009), children identified with language impairments (S. L. Gillam & Gillam, 2016; S. L. Gillam et al., 2018), and children with autism spectrum disorder (S. Gillam et al., 2015) show that SKILL is associated with improved narrative language outcomes across subgroups of children. Two preliminary studies assessed the feasibility and promise of the SKILL

program with bilingual students in the elementary grades with developing English language proficiency. Jensen (2009) found that bilingual students' narratives in both English and Spanish increased in complexity after receiving intervention with an early version of the SKILL program. In a study of the current version of the SKILL program, S. L. Gillam et al. (2018) found that the SKILL intervention was associated with substantial improvements in narrative language for bilingual students. Results of these single-group, pre–post design studies supported the feasibility of the program for use with bilingual children. An experimental design with a larger sample of bilingual students that also measures the effects of SKILL on reading comprehension was necessary.

This Study

This study sought to expand our understanding of evidence-based practices for bilingual students by addressing two questions using data from S. L. Gillam et al.'s (2023) RTC. First, do the effects of SKILL vary between bilingual and monolingual students on English narrative language and reading comprehension? We were interested in evaluating the impact of SKILL for EBs for a few reasons: (a) Narrative language represents an area of vulnerability because it requires children to integrate semantics, morphology, syntax, and pragmatics and to organize this information in a logical sequence (Squires et al., 2014; Uccelli & Páez, 2007), and (b) many EBs have difficulty meeting grade-level expectations for narrative language and text comprehension (NAEP, 2022), and yet (c) relatively few narrative language intervention studies have examined the effects for EBs (e.g., R. D. Miller et al., 2018), and the large-scale RCTs previously reported did not include a large number of EBs (e.g., LARRC et al., 2019; Petersen et al., 2022). Although examining approaches to supporting children's home language in tandem with their second language is critical, the shortage of bilingual teachers and SLPs (e.g., ASHA, 2018) underscores the need to assess the effects of interventions that are provided in English, such as SKILL. Another reason we were motivated to examine primary language status was that prior research has shown that EB status moderates the effects of academic intervention in some cases (e.g., Llosa et al., 2016) but not in others (e.g., Vaughn et al., 2017). However, this question has not yet been addressed in the context of a narrative intervention for school-age children with English language and literacy difficulties. We sought to explore this question related to differential efficacy based on language status using moderation because it enables us to understand the extent to which the effects of SKILL are universal (Kenny, 2018). This approach has advantages over other approaches to analyzing subgroups of students, such as conducting multiple separate significance tests, because it is less susceptible to Type I error (Wang & Ware, 2013).

Our second research question focused on better understanding the relation between students' initial performance and their response to instruction: To what extent does initial narrative language performance in English differentially predict response to intervention for monolinguals and bilinguals? We were interested in whether initial narrative language abilities in English moderated the effects of intervention for bilingual students due to earlier research that shows their response to intervention varied based on initial language and reading skill levels (e.g., Lovett et al., 2008; Stuebing et al., 2015), including among students who are bilingual (Vaughn et al., 2019). For example, Lovett et al. (2008) conducted an RCT to explore the impact of a phonologically based reading intervention for struggling readers from different primary language backgrounds (i.e., monolingual English, English language learner [ELL]) who also demonstrated low oral language abilities in English. Struggling readers who participated in the experimental reading intervention outperformed those in the control condition. Of particular interest was the finding that initial oral language status was a strong predictor of reading growth. The students who presented lower language abilities (e.g., language impairments) were observed to benefit the most from the reading instruction. Also of interest was the finding that primary language status (monolingual English, ELL) did not impact outcomes. Together, these questions will contribute to our knowledge about the efficacy of SKILL for bilingual students with developing English language proficiency and help address gaps in the present research that are particularly important for SLPs and, potentially, general and special education teachers who provide Tier 2 interventions.

Method

This article uses data from a multisite, multicohort RCT (S. L. Gillam et al., 2023) collected to evaluate the efficacy of SKILL. After obtaining institutional review board approval, researchers implemented the SKILL efficacy trial with three cohorts of students (one cohort per year for three consecutive years). Each year, the multistep participant-screening process began in November, pretesting occurred in December, instruction lasted from January to March, and posttesting occurred in April. The cohorts did not vary in their procedures except for Cohort 3, which was impacted by the COVID-19 pandemic (additional information provided below). The research team systematically identified elementary students in Grades 1–4 with English language and literacy difficulties. A total of 14 school sites from seven school districts, located in a mix of urban, near-urban, and rural school districts in one state

each of Southwestern and Western United States, participated in the study.

Setting and Participants

The research team used a multiple-screening process to identify students in Grades 1–4 at risk for difficulties in narrative language and reading comprehension in English for participation. Similar to other studies of at-risk children (e.g., Coyne et al., 2013), our sample included students who performed at or below the bottom tertile (33rd percentile) on a reading comprehension measure (GMRT-RC; MacGinitie et al., 2001) and a narrative language measure (Test of Narrative Language–Second Edition [TNL-2]; R. B. Gillam & Pearson, 2017), including students with disabilities. A total of 356 students over 3 years met the qualifying screening criteria and received parent consent for participation. We randomly assigned these students, blocked by classroom and stratified by grade level, to receive the SKILL intervention or a BAU instruction. Thus, this study represents a nested design, with students partially nested in tutors, tutors nested in teachers, and teachers nested in schools.

One student was excluded from this study because bilingual status was unknown. Hence, the sample for this study included 355 students. The ethnic composition included students who were Latino/a (56.7%), White (31.6%), African American (4.5%), and Asian (1.4%). Thirty-seven percent of students had a previously identified disability, and 45% received free or reduced lunch. The most common disability categories among participating students were speech-language impairments (19%) and learning disabilities (16%). The participating sample included a large number of bilingual students (n = 148; 41% of the total sample), as identified by parents who reported that their child primarily spoke a language other than English at home. Of these bilingual students, 139 (94%) spoke Spanish as their primary language at home (other primary languages included Arabic, French, Khmer, Korean [two students], Lingala, Polish, Somali, and Swahili). None of the bilingual students were newcomers in their first year of English language instruction. Such students were excluded from the study because they may appear at risk for English language and reading comprehension difficulties due to their limited opportunities to engage in English language instruction. Of the bilingual students, parents and guardians reported that 51% of these students were proficient in their home language (e.g., could speak in full conversations). The other family members (49%) reported that their child had limited home language ability (e.g., spoke in phrases, but were not fully conversational). A majority of families (62%) reported that they spoke a language other than English with their child at least 1 hr a day at home. Only 21% of all participating children had received formal instruction in their home language (e.g., received language tutoring outside of school or dual language instruction at school). Taken together, there was substantial heterogeneity within the sample of EBs, both in their home language proficiency—their daily home language use—and in the instruction students received in their home language. However, this reflects the profound heterogeneity that exists within the population of EBs (Mancilla-Martinez & Lesaux, 2011).

Table 1 presents demographic data on the participants by language status (i.e., bilingual or monolingual). Most critical to this study, there were no significant differences between bilingual and monolingual students on demographic variables (i.e., grade level, gender, age) except for the proportion of students identified with disabilities ($\chi 2$ = 11.89; p = .00). A greater number of monolingual students were identified with disabilities compared to bilingual students. Prior research suggests that language-minority children are less likely to be identified as having learning disabilities and speech-language impairments than their language-majority peers (Morgan et al., 2015). To exclude the possibility that bilingual and monolingual students' differential response to intervention was a function of special education status, we examined the three-way interaction between language status, condition, and special education status. The interaction terms did not differ from 0 for any of the outcomes, suggesting that the two-way interaction between condition and bilingual status was not affected by special education status (p values ranged from .11 to .95).

Of the 355 students included in this study, 52 students attrited over the duration of the study because their family moved during the course of the intervention (n = 13) or did not respond to the request to complete posttesting after COVID-19 closed schools (n = 39). We report rates of attrition as the ratio of randomized units to the sample at posttest. Sample-wide attrition was 15%. The difference in attrition rates for SKILL (.14) and BAU (.16) was .02. The combination of these rates of attrition—overall and differential attrition—represents low threats to the internal validity of the study based on standards recommended by the What Works Clearinghouse (2020). Also critical to this study, the rates of attrition for bilinguals (.14) and monolinguals (.15) were similar.

Description of the SKILL Intervention

The SKILL program addresses oral narrative language proficiency through instruction targeting macrostructural (story elements and the causal relationship between them) and microstructural (complex linguistic structures including adverbial and relative clauses) aspects of narratives in the context of authentic narrative

Table 1. Student demographics.

	SKILL (*N* = 185)				BAU (*N* = 170)			
	Monolingual (*n* = 102)		Bilingual (*n* = 83)		Monolingual (*n* = 105)		Bilingual (*n* = 65)	
Demographic variable	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Gender								
Male	64	62.7	43	51.8	49	46.7	37	56.9
Female	38	37.3	40	48.2	56	53.3	28	43.1
Ethnicity								
African American	6	5.9	3	3.6	5	4.8	2	3.1
Asian	0	0	3	3.6	2	1.9	0	0
Latino/a	34	33.3	71	85.5	40	38.1	58	89.3
White	56	54.9	3	3.6	52	49.5	1	1.6
Two or more	1	1	1	1.2	4	3.8	3	4.4
Other	2	2	2	2.4	1	1	0	0
Missing	3	2.9	0	0	1	1	1	2
Students with disabilities	52	51	21	25.3	42	40	20	30.8
Age in years (*M*, *SD*)	8.53	1.12	8.53	1.16	8.45	1.07	8.67	1.21
Grade level								
1	18	17.6	13	15.7	18	17.1	12	18.5
2	25	24.5	22	26.5	31	29.5	13	20
3	33	32.4	25	30.1	36	34.3	15	23.1
4	26	25.5	23	27.7	20	19	25	38.5

Note. SKILL = Supporting Knowledge in Language and Literacy; BAU = business as usual.

discourse. Each lesson is fully manualized and includes objectives, instructions, materials, video examples, and lesson plans, with teacher scripting for each instructional session. The SKILL program includes 37 lessons written in English and organized into three instructional phases: (I) teaching story structure and causal language, (II) teaching strategies for creating a situation model (overall theme of the story), and (III) teaching strategies for integration into long-term memory. Phase I contains 18 lessons that provide students with an understanding of the main story elements, including characters, setting, initiating event, internal response, plans, actions, and consequences in the context of a wordless picture story. In Unit 1 lessons, tutors teach students story elements and a representative icon, which is situated on a sequenced storyboard that later serves as a graphic organizer. The story grammar elements are taught explicitly to students, and then, students are provided opportunities to identify the elements in the context of a story. Also in Phase I, tutors help students build their knowledge of story elements by retelling a story, initially with teacher support. Finally, tutors work with their small group of students to develop their own story using their storyboard graphic organizer.

In Phase II, students are taught how to create longer stories with more complex temporal and causal relationships and how to make stories more interesting through the inclusion of more expressive verbs (e.g., "yelled," "growled," "whispered"; "thought," "decided," "promised") and character dialogue in eight lessons. For example, tutors teach students what elaboration and dialogue mean and how to incorporate these in their stories. Just like in Phase I, students work on developing more complex stories by listening to stories and creating their own stories.

Phase III contains 12 lessons to give students multiple opportunities to retell, create, tell, edit, and revise their own spontaneously generated stories with and without icon and graphic organizer support. The major focus of Phase III is to provide students with opportunities to develop independence in their understanding and use of narrative structure (story elements and causal framework) and complex oral language relevant to the production of coherent, organized, and memorable stories (connecting terms, causal language, mental state terms). Lessons allow students opportunities to practice actively storing and retrieving information from long-term memory. Further information about SKILL is presented in other articles (S. L. Gillam & Gillam, 2016; S. L. Gillam et al., 2023).

Intervention and BAU Instruction Procedures

Participants randomized to the treatment condition participated in thirty-six 30-min SKILL instructional sessions as a supplemental intervention provided by a member of the research staff. These sessions occurred in small groups ranging in size from two to four students and took place in other classrooms outside of the general education

classroom. Students were placed into small groups with their grade-level peers (i.e., first graders were with other first graders) based on when their teachers made them available. When more than four students were available at a specific time, we randomly assigned students to treatment groups. Instruction occurred 3–5 days per week and ranged from 9 to 12 weeks in duration. Students participated in instruction outside of the English language arts (ELA) block (that time in which teachers provide ELA instruction, typically using materials from their adopted ELA curriculum). This ensured that all students (BAU and treatment students) received their ELA instruction from their classroom teachers. There was considerable variation in the instruction students randomized to the BAU condition received during the time treatment students were provided SKILL instruction. Most students received small-group reading or math intervention from their classroom teacher or an aide. Observations of classroom teachers revealed that there was minimal overlap between classroom teachers' instruction and SKILL instruction because classroom teachers focused little instructional time on narrative language learning (Hall et al., 2021).

Interventionists and Treatment Fidelity

Twenty-four tutors (92% women) hired and trained by the research staff implemented instruction across the three cohorts to students randomized to the SKILL treatment condition. Of the 24 tutors, 19 held a credential in teaching or speech-language pathology. The five remaining tutors were enrolled in a graduate program, studying communication disorders or special education, and had extensive experience tutoring in authentic educational contexts or past research projects. Each year, tutors received full-day training on the SKILL program from the principal investigators (PIs) who also led the development of the SKILL program. During the training, the PI described and modeled lessons in each phase. Video recordings of gold standard instruction for each lesson and hand-selected video examples from prior years were shown during training and were made available to the tutors throughout the project. Tutors were given opportunities to practice delivering the lessons, ask questions, and receive feedback during training and throughout the delivery of instruction. After the training, tutors were asked to teach a mock session to an experienced field supervisor with greater than 90% fidelity before working with students.

Throughout the duration of the intervention, the research team video- and audio-recorded all instructional sessions. Every session was rated for fidelity by a research assistant whose primary responsibility was fidelity monitoring. The PI or a project coordinator met immediately with any tutor who scored below 80% overall fidelity for an instructional session and provided corrective feedback. Weekly ongoing trainings were held with all tutors to address concerns and prepare for upcoming sessions.

The evidence to demonstrate that the instruction was implemented with high fidelity included (a) the completion of an instruction observation checklist (IOC) designed for each lesson and (b) fidelity follow-up meetings including debriefing procedures. The IOC assessed the interventionists' adherence to the essential components of each lesson of SKILL program, which including following the sequence of the scripted lessons and utilizing language facilitation techniques, key words, and the appropriate instructional materials (e.g., icons, storyboards, self-scoring rubrics). Overall, treatment fidelity was very high across all instructional tutors (M = 96%). Average fidelity was consistently high across interventionists (range: 85%–100%), research sites (M = 96% at both sites), and cohorts (range: 90%–98%). Further information about treatment fidelity is presented in the article reporting the main effects of the study (Gillam et al, 2023).

Assessment Procedures and Measures

All assessments were administered in English by trained testers who were blind to the students' study condition. These assessments occurred at the participating campuses in separate classrooms. Pretest assessments occurred within 2 weeks of the treatment instruction beginning, and posttest assessments occurred within 2 weeks of the conclusion of instruction. Senior researchers responsible for data collection hired and trained all assessment administrators (a separate team from the intervention unit). All assessment administrators were members of the research team (either graduate students or assessment staff members), with experience administering assessments following standardized protocols. Before each testing period, testers received full-day training on how to deliver the assessments with fidelity and reliability. In order to be cleared for testing, each assessment administrator demonstrated that they could administer the assessment with fidelity and reliability during a mock session with a senior researcher. All assessments were double-scored before data entry.

TNL

The TNL-2 (R. B. Gillam & Pearson, 2017) is a norm-referenced tool designed to be a stable measure of narrative comprehension and production abilities normed on students between the ages of 4 and 15 years. There are three narrative formats in the TNL-2. In the first context, students are told a scripted story about two children who go to a fast-food restaurant with their mother. After students answer a series of literal and inferential comprehension questions about the story, they are directed to retell it. In the second context, students are shown a series of

five sequenced pictures about a commonly experienced event (completing a homework project and taking it back to school). After listening to a one-episode story that corresponds with the pictures, they are presented with a series of literal and inferential questions about the story. Then, they are asked to generate their own story based on a sequenced picture prompt about a boy who is late for school. In the final and most challenging context, students are shown a single-scene picture prompt about a dragon guarding a treasure chest and are asked to listen to a two-episode story about the picture and then answer literal and inferential questions about it. Finally, students are shown a new single-scene picture prompt about aliens landing in a park and are asked to generate a story about it. The story productions are scored according to critical content, language complexity, and coherence. The TNL-2 takes between 15 and 25 min to administer. Test–retest reliability on the TNL-2 is .93. R. B. Gillam et al. (2013) reported evidence to support the validity of the TNL among bilingual children receiving instruction in English for a year or more.

Monitoring Indicators of Scholarly Language

Students were asked to create their own stories orally and in written form before and immediately after those in the experimental group participated in the SKILL intervention program. Students were shown a picture prompt depicting an event (jeep in the desert, plane landing) and asked to tell (write) the best story they could. The icons used in the instructional program were placed on the table in front of the students, but no explanation of how to use them was provided. Oral and written stories composed by students were transcribed and scored using the Monitoring Indicators of Scholarly Language (MISL; S. L. Gillam et al., 2017) rubric, which is a progress-monitoring tool that was designed to capture changes in the use of macrostructure and microstructure that were the focus of instructional lessons. Seven macrostructure elements (character, setting, initiating event, internal response, plan, action, and consequence) are assessed and scored on a 4-point scale (0–3). Similarly, seven microstructure elements (coordinating conjunctions, subordinating conjunctions, metacognitive verbs, metalinguistic verbs, elaborated noun phrases, grammaticality, and tense) are assessed on a 4-point scale to evaluate the types of words the students use to connect phrases and to express meaning in created stories. The MISL subtest scores range from 0 to 42. The MISL measures of narrative macrostructure and microstructure elements were closely aligned with changes in story production that should occur after SKILL intervention. Internal consistency reliability for the MISL is .79. Members of the assessment team calculated interrater agreement for 20% of all the MISL samples that were scored independently and found that it was high on both oral (95.7%) and written (93.2%) narratives.

GMRT

The GMRT-RC (MacGinitie et al., 2001) is a group-administered, timed assessment designed to access a student's reading comprehension abilities. Students silently read expository and narrative passages that range from three to 15 sentences in length and respond to multiple-choice comprehension questions. For the relevant grade levels, internal consistency reliability for the GMRT-RC ranges from .91 to .93, with alternate-form reliability ranging from .80 to .87 (MacGinitie et al., 2001).

Test of Silent Reading Efficiency and Comprehension

The Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner et al., 2010) is a timed reading measure that assesses a student's reading fluency and comprehension performance. The TOSREC requires students to read a list of sentences silently and then respond if the given sentence is true or false by circling "yes" or "no." Students have 3 min to respond to as many sentences as they can. The TOSREC reports an alternate-form reliability range of .86–.95, with a test–retest (2 months) alternate-form reliability range of .81–.87 (Wagner et al., 2010). For Grades 1–5, the TOSREC has a strong concurrent and predictive correlation with oral reading fluency, with an average coefficient of .73 (Wagner et al., 2010). Cronbach's α for the TOSREC is reported as .97.

Test of Word Reading Efficiency–Second Edition Sight Word Efficiency Subtest

The Test of Word Reading Efficiency–Second Edition (TOWRE-2; Torgesen et al., 2012) Sight Word Efficiency (SWE) subtest consists of real words the students are asked to read that increase in difficulty as the student advances. The SWE subtest is scored by how many correct words the student reads in those 45 s. Average alternative-form coefficients for the TOWRE-2 were .91 on the SWE subtest. Test–retest reliabilities ranged from .89 to .93 with interrater reliability being .99 (Torgesen et al., 2012). We report standard scores (M = 100, SD = 15) based on grade level.

Impact of COVID-19 on Study Procedures

The third cohort occurred during the 2019–2020 school year and was impacted by the COVID-19 pandemic. In March of 2020, students randomized to treatment had received about two thirds of their instructional sessions in person (about 24 of 36 sessions) when schools were closed at both research sites as a result of the COVID-19 pandemic for the rest of the school year. Our research team worked directly with families to complete the remaining instructional sessions with treatment students and, posttest, with the treatment and BAU students

individually via videoconference. By mailing materials to families, we were able to follow the same instruction and assessment procedures used during in-person instruction while schools were closed except for interactions that occurred over videoconference. Treatment fidelity data suggest that the intervention was implemented with high fidelity, as adherence was measured at greater than 96% during the videoconferenced lessons. Assessment procedures aligned with recommendations and guidelines for the remote assessment of children, including using a videoconferencing platform to administer tests, ensuring children have access to a strong Wi-Fi connection, minimizing noise distractions, and working closely with families to ensure that they can assist with scheduling testing sessions and troubleshooting technology issues but refrain from providing any academic supports or feedback during testing (Garrisi et al., 2020). As with previous testing periods, each assessment administrator demonstrated that they could administer the assessment with fidelity and reliability during a mock session with a senior researcher. An evaluation of the psychometric properties of the online-administered format of the TNL-2 (R. B. Gillam & Pearson, 2017) revealed similar internal consistency reliability, scorer reliability, and measurement invariance for online and in-person administration (Magimairaj et al., 2022). In the experimental context, there were no significant differences by cohort, which indicates that the effect did not vary for Cohort 3. Further information about study procedures in response to COVID-19 is provided in the work of S. L. Gillam et al. (2023). It is also important to consider that this study derives from a large-scale RCT. RCTs are considered the gold standard method because the processes used in the implementation of an RCT minimize threats to validity (Akobeng, 2005). Although it is possible there are unmeasured factors related to COVID-19 that influenced study outcomes for the third cohort of students who experienced the pandemic, there are no reasons to think these factors would influence students randomized to treatment or control differently.

Data Analysis Plan

In the study from which these data derive (S. L. Gillam et al., 2023), students were randomized to one of two conditions—SKILL treatment or BAU. We used multilevel models to examine SKILL's main treatment effect on student outcomes. Those results, reported elsewhere (S. L. Gillam et al., 2023), are experimental. We consider this study, which focuses on the conditional effect of language status, to be quasi-experimental because language status was not included as a stratum when randomizing to condition. For that reason, we model language status as a potential moderator of SKILL's effect. We initially estimated a four-level model, with random intercepts at all four levels. With the exception of MISL Oral Narrative (interclass correlation coefficient = .11), there was no variance at the tutor level, so it was removed from all the models except the models for the MISL Oral Narrative–dependent measure. In fitting the conditional models for all outcome variables expect the GMRT, teacher-level variance became close to 0, so we dropped it from the models. School-level variance remained significant for all the variables expect the GMRT.

The first research question considers the extent to which SKILL's effect differs on average for bilingual students and students whose first language is English. To answer this research question, we expand the models that estimate the main effect of the SKILL intervention to include a conditional effect representing language status' potential moderation of the treatment effect (i.e., its interaction with the treatment condition). The reduced-form equation for TNL Comprehension, TNL Production, MISL Written Narrative, and the TOSREC is as follows:

$$\begin{aligned} Y_{ik} = {} & \gamma_{00} + \gamma_{10}(Pretest)_{ik} + \gamma_{20}(SKILL)_{ik} + \gamma_{30}(Female)_{ik} \\ & + \gamma_{40}(SWD)_{ik} + \gamma_{50}(Language)_{ik} + \gamma_{60}(Grade\ level)_{ik} \\ & + \gamma_{70}(SKILL \times Language)_{ik} + u_{0k} + e_{ik}. \end{aligned} \quad (1)$$

The equations for MISL Oral Narrative (Equation 2) and GMRT Comprehension (Equation 3) are as follows:

$$\begin{aligned} Y_{itk} = {} & \gamma_{000} + \gamma_{100}(Pretest)_{itk} + \gamma_{200}(SKILL)_{itk} \\ & + \gamma_{300}(Female)_{itk} + \gamma_{400}(SWD)_{itk} + \gamma_{500}(Language)_{itk} \\ & + \gamma_{600}(Grade\ level)_{itk} + \gamma_{700}(SKILL \times Language)_{itk} u_{00k} \\ & + u_{0tk} + e_{itk} \end{aligned} \quad (2)$$

$$\begin{aligned} Y_{ij} = {} & \gamma_{00} + \gamma_{10}(Pretest)_{ij} + \gamma_{20}(SKILL)_{ij} + \gamma_{30}(Female)_{ij} \\ & + \gamma_{40}(SWD)_{ij} + \gamma_{50}(Language)_{ij} + \gamma_{60}(Grade\ level)_{ij} \\ & + \gamma_{50}(SKILL \times Language)_{ij} u_{0j} + e_{ij}. \end{aligned} \quad (3)$$

Here, subscript "i" represents students, subscript "t" represents tutors, subscript "j" represents teachers, and subscript "k" represents schools. Parameters γ_{000} and γ_{00} are the student-level mean outcomes for each measure in the study; $Pretest_{ik}$, $Pretest_{itk}$, $Pretest_{ij}$, and $Pretest_{itj}$ are student-level pretest scores for each outcome centered around its grand mean (Enders & Tofighi, 2007); $SKILL_{ik}$, $SKILL_{itk}$, and $SKILL_{ij}$ are student-level, dummy-coded variables representing condition, where SKILL intervention is coded as 1 and BAU is coded as 0; $Female_{ik}$, $Female_{itk}$, and $Female_{ij}$ are the students' gender, with male coded as 0 and female coded as 1; SWD_{ik}, SWD_{itk}, and SWD_{ij} refer to student with disability status, with the non-SWD group coded as 0 and the SWD group coded as

1; $Language_{ik}$, $Language_{itk}$, and $Language_{ij}$ are language status, with monolinguals coded as 0 and bilinguals coded as 1; $Grade\ level_{ik}$, $Grade\ level_{itk}$, and $Grade\ level_{ij}$ are the students' grade level, with the first and second graders coded as 1 and the third and fourth graders coded as 0; $SKILL \times Language_{ik}$, $SKILL \times Language_{itk}$, and $SKILL \times Language_{ij}$ represent the interactions between language status and treatment assignment; and residuals are Level 1, Level 2, and Level 3 random effects, respectively.

The second research question examines narrative language performance at pretest and its potential moderating effect on SKILL's impact for bilinguals and monolinguals. Scores from the Comprehension and Production subtests of the TNL were added to Equations 1–3 before estimating all two-way interactions and the three-way interaction between TNL, language status, and condition. TNL Time 1 scores were grand mean centered.

We applied the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995) to control for false discovery rates (Type 1 error) with multiple comparisons (What Works Clearinghouse, 2020). All analyses were run with the lme4 package in R (Bates et al., 2015). Two-way interactions were decomposed and contrasts were computed using the emmeans package in R (Lenth, 2020) in R. We computed Hedges's g as a covariate-adjusted mean difference divided by the unadjusted pooled within-group standard deviation (What Works Clearinghouse, 2020).

Results

Research Question 1: To What Extent Do the Effects of SKILL Vary Between Bilingual and Monolingual Students on English Narrative Language and Reading Outcomes?

Table 2 summarizes observed pre- and posttest means and standard deviations for the two intervention conditions and for bilingual and monolingual students. Tables 3 and 4 present the results of the moderation analysis in three models. Model 1 in Tables 3 and 4 addresses the first research question and looks at the extent to which the effect of SKILL varies between bilingual and monolingual students. The second and third models in Tables 3 and 4 address the second research question. Models 2 and 3 test whether students' narrative language performance at pretest as measured by TNL Comprehension (Model 2) and TNL Production (Model 3) in English predicted

Table 2. Pretest and posttest means and standard deviations for outcome measures.

	Monolingual			Bilingual			Monolingual			Bilingual		
	Pretest			Posttest			Pretest			Posttest		
Outcome variable	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
Language outcomes												
TNL Comprehension												
BAU	105	6.54	1.99	65	5.75	2.55	93	7.95	2.53	53	6.49	2.45
SKILL	102	6.45	2.13	83	6.07	2.30	88	7.90	2.22	71	7.76	2.22
TNL Production												
BAU	105	6.32	2.00	65	6.25	2.33	92	7.25	2.36	53	7.19	2.14
SKILL	102	6.54	1.94	83	6.46	1.84	87	8.32	2.41	71	7.87	2.27
MISL Written Narrative												
BAU	105	7.77	5.37	65	8.94	6.13	96	8.24	5.43	54	9.65	6.17
SKILL	102	7.74	4.90	83	8.69	5.49	89	10.99	7.02	73	10.48	6.76
MISL Oral Narrative												
BAU	105	10.14	5.74	65	11.52	5.91	95	11.35	5.80	54	10.65	6.69
SKILL	102	10.28	5.63	83	11.06	7.05	89	15.73	7.29	73	14.49	7.41
Reading outcomes												
GMRT												
BAU	105	82.32	9.62	65	83.64	8.51	93	87.10	12.16	54	85.35	9.72
SKILL	102	82.47	8.83	83	81.76	9.92	86	86.81	11.46	71	83.98	11.61
TOSREC												
BAU	105	13.86	8.48	65	12.46	10.59	93	18.55	8.91	54	16.24	9.09
SKILL	102	14.15	8.85	83	13.07	9.15	86	17.41	8.57	72	15.08	8.96

Note. TNL = Test of Narrative Language; BAU = business as usual; SKILL = Supporting Knowledge in Language and Literacy; MISL = Monitoring Indicators of Scholarly Language; GMRT = Gates–MacGinitie Reading Test; TOSREC = Test of Silent Reading Efficiency and Comprehension.

Table 3. Results of the moderation analysis for language outcomes.

Fixed effects	Model 1			Model 2			Model 3		
	Est.	*SE*	*p*	Est.	*SE*	*p*	Est.	*SE*	*p*
TNL Comprehension									
Intercept	8.01	0.32	< .001	7.97	0.32	< .001	8.04	0.31	< .001
Pretest	0.54	0.05	< .001	0.65	0.1	< .001	0.52	0.06	< .001
SKILL	0	0.3	.99	0.04	0.3	.91	−0.05	0.3	.88
Female	0.38	0.24	.11	0.39	0.24	.1	0.33	0.24	.16
SWD	−0.33	0.25	.19	−0.3	0.25	.23	−0.23	0.25	.35
Grade level	−0.37	0.24	.12	−0.39	0.24	.1	−0.46	0.24	.05
Language	−1.1	0.35	< .001	−1.04	0.36	< .001	−1.11	0.35	< .001
SKILL × Language	**1.15**	**0.48**	**.02**	1.06	0.48	.03	1.21	0.47	.01
Moderator							0.25	0.1	.02
SKILL × Moderator				−0.14	0.14	.33	−0.12	0.15	.45
Language × Moderator				−0.04	0.15	.8	−0.09	0.16	.55
SKILL × Language × Moderator				−0.06	0.21	.79	−0.16	0.23	.49
TNL Production									
Intercept	7.54	0.34	< .001	7.51	0.33	< .001	7.55	0.34	< .001
Pretest	0.35	0.06	< .001	0.27	0.06	< .001	0.38	0.11	< .001
SKILL	1.05	0.32	< .001	1.1	0.31	< .001	1.06	0.32	< .001
Female	0.59	0.25	.02	0.58	0.24	.02	0.59	0.25	.02
SWD	−0.73	0.26	.01	−0.79	0.25	< .001	−0.74	0.26	.01
Grade level	−0.24	0.25	.33	−0.3	0.24	.21	−0.24	0.25	.34
Language	−0.1	0.37	.79	0.08	0.36	.82	−0.11	0.37	.77
SKILL × Language	−0.51	0.5	.3	−0.62	0.48	.2	−0.52	0.5	.3
Moderator				0.38	0.1	< .001			
SKILL × Moderator				−0.11	0.14	.45	−0.06	0.16	.69
Language × Moderator				−0.1	0.16	.54	−0.1	0.16	.55
SKILL × Language × Moderator				−0.02	0.21	.94	0.25	0.25	.31
MISL Written Narrative									
Intercept	9.05	0.71	< .001	9.05	0.71	< .001	9.04	0.72	< .001
Pretest	0.66	0.06	< .001	0.65	0.06	< .001	0.65	0.06	< .001
SKILL	2.98	0.73	< .001	2.89	0.74	< .001	3.02	0.74	< .001
Female	0.81	0.57	.16	0.88	0.58	.13	0.83	0.58	.16
SWD	−0.06	0.59	.91	0	0.6	.99	−0.08	0.6	.89
Grade level	−1.99	0.63	< .001	−1.95	0.63	< .001	−1.98	0.64	< .001
Language	0.46	0.84	.58	0.48	0.86	.57	0.49	0.84	.56
SKILL × Language	−1.98	1.15	.08	−2.04	1.16	.08	−2.06	1.16	.08
Moderator				−0.34	0.25	.17	0.03	0.26	.89
SKILL × Moderator				0.34	0.35	.34	−0.25	0.37	.51
Language × Moderator				0.53	0.37	.16	0.13	0.38	.73
SKILL × Language × Moderator				−0.7	0.51	.17	0.18	0.57	.75

(*table continues*)

bilingual and monolingual students' differential response to instruction.

Language Outcomes

TNL comprehension. The two-way interaction of SKILL and language status was positive and statistically significant (γ_{40} = 1.15, *SE* = 0.48, *p* = .02). This means that the intervention affected narrative comprehension in bilinguals and monolinguals differently, with the effect of the intervention being greater for bilingual students than for monolingual students. As shown in Table 5, the effect of treatment on narrative comprehension was 0.00 (95% CI [−0.29, 0.29]) for monolinguals but 0.50 (95% CI [0.13, 0.86]) for bilinguals.

TNL production. The two-way interaction of SKILL and language status did not differ statistically from 0 (γ_{40} = −0.51, *SE* = 0.50, *p* = .30), indicating that the intervention's effect did not differ for the two groups. The treatment effect size was 0.44 (95% CI [0.14, 0.74]) for the

Table 3. *(Continued).*

Fixed effects	Model 1			Model 2			Model 3		
	Est.	***SE***	***p***	**Est.**	***SE***	***p***	**Est.**	***SE***	***p***
MISL Oral Narrative									
Intercept	12.71	0.85	< .001	12.64	0.85	< .001	12.72	0.84	< .001
Pretest	0.51	0.05	< .001	0.51	0.06	< .001	0.51	0.06	< .001
SKILL	4.41	0.91	< .001	4.39	0.91	< .001	4.39	0.9	< .001
Female	1.09	0.64	.09	1.15	0.65	.08	1.07	0.65	.1
SWD	−0.24	0.67	.72	−0.16	0.68	.82	−0.07	0.67	.92
Grade level	−3.12	0.67	< .001	−3.1	0.68	< .001	−3.22	0.69	< .001
Language	−1.26	0.97	.19	−1.11	0.99	.26	−1.29	0.96	.18
SKILL × Language	−0.61	1.31	.64	−0.84	1.33	.53	−0.51	1.31	.7
Moderator				−0.04	0.28	.9	0.43	0.29	.14
SKILL × Moderator				0.13	0.39	.73	−0.45	0.42	.28
Language × Moderator				0.33	0.42	.43	−0.2	0.43	.65
SKILL × Language × Moderator				−0.75	0.57	.19	−0.42	0.64	.52
Random effects	**Var.**	**ICC**		**Var.**	**ICC**		**Var.**	**ICC**	
TNL Comprehension									
Student level	3.89	.94		3.89	.94		3.83	.95	
School level	0.23	.06		0.23	.06		0.2	.05	
TNL Production									
Student level	4.22	.93		3.93	.93		4.25	.93	
School level	0.3	.07		0.28	.07		0.3	.07	
MISL Written Narrative									
Student level	23.48	1		23.62	1		23.7	.99	
School level	0.12	0		0.06	0		0.14	.01	
MISL Oral Narrative									
Student level	26.33	.84		26.72	.85		26.49	.86	
Tutor level	3.89	.12		3.63	.12		3.59	.12	
School level	0.99	.03		0.92	.03		0.75	.02	

Note. Bolded values represent statistically significant results related to our research questions. "Moderator" refers to TNL Comprehension in Model 2 and TNL Production in Model 3. Est. = Estimate; *SE* = standard error; TNL = Test of Narrative Language; SKILL = Supporting Knowledge in Language and Literacy; MISL = Monitoring Indicators of Scholarly Language; Var. = variance; ICC = interclass correlation coefficient.

monolingual group and 0.24 (95% CI [−0.11, 0.60]) for the bilingual group.

MISL oral narrative. The positive effect of the SKILL treatment on narrative production did not differ for bilingual and monolingual students, as indicated by the nonsignificant interaction term (γ_{40} = −0.61, *SE* = 1.31, *p* = .64). The effect sizes for SKILL in the monolingual and bilingual groups were 0.67 (95% CI [0.37, 0.97]) and 0.53 (95% CI [0.18, 0.89]), respectively.

MISL written narrative. The positive effects of the SKILL treatment did not differ for bilingual and monolingual students, as indicated by the nonstatistically significant interaction term (γ_{40} = −1.98, *SE* = 1.15, *p* = .08). This means that the effect of SKILL on students' abilities to write narratives about a picture prompt did not vary based on language status. The effect sizes for SKILL in the monolingual and bilingual groups were 0.48 (95% CI [0.18, 0.77]) and 0.15 (95% CI [−0.20, 0.51]), respectively.

Generalized Outcomes: Reading

GMRT. Reading outcomes are summarized in Table 4. According to results on the GMRT, the simple main effect for treatment on reading comprehension performance did not differ from 0 (γ_{10} = −0.13, *SE* = 1.53, *p* = .93), on average, and the treatment's effect did not differ for bilingual and monolingual students. The effect size in the monolingual group was −0.02 (95% CI [−0.31, 0.28]). In the bilingual group, the SKILL effect was −0.03 (95% CI [−0.33, 0.38]). On the basis of our previous finding that the effect of SKILL on the GMRT varied as a function of students' grade level (S. L. Gillam et al., 2023), we tested the three-way interaction between bilingual status, SKILL, and grade level and it was not statistically significant (*p* = .39).

TOSREC. On the TOSREC, students' performance in the SKILL condition (γ_{10} = −1.34, *SE* = 1.09, *p* = .22) did not differ, on average, from that of students in the BAU condition, and the treatment's effect on reading efficiency

Table 4. Results of the moderation analysis for reading outcomes.

	Model 1			Model 2			Model 3		
Fixed effects	**Est.**	***SE***	***p***	**Est.**	***SE***	***p***	**Est.**	***SE***	***p***
GMRT									
Intercept	86.7	1.59	< .001	86.92	1.58	< .001	86.74	1.59	< .001
Pretest	0.34	0.07	< .001	0.31	0.07	< .001	0.34	0.07	< .001
SKILL	−0.13	1.53	.93	−0.16	1.52	.92	−0.29	1.54	.85
Female	0.44	1.21	.72	0.21	1.2	.86	0.34	1.22	.78
SWD	−2.66	1.28	.04	−2.83	1.28	.03	−2.41	1.29	.06
Grade level	4.5	1.42	< .001	4.13	1.41	< .001	4.18	1.44	< .001
Language	−3.4	1.81	.06	−2.73	1.81	.13	−3.24	1.82	.08
SKILL × Language	−0.11	2.44	.96	−0.56	2.43	.82	−0.01	2.45	1
Moderator				0.71	0.53	.18	0.29	0.55	.6
SKILL × Moderator				0.15	0.73	.84	0.49	0.8	.54
Language × Moderator				0.72	0.78	.36	0.35	0.81	.67
SKILL × Language × Moderator				−1.04	1.06	.33	−1.72	1.21	.16
TOSREC									
Intercept	18.72	1.11	< .001	18.76	1.11	< .001	18.77	1.1	< .001
Pretest	0.59	0.06	< .001	0.58	0.06	< .001	0.58	0.06	< .001
SKILL	−1.34	1.09	.22	−1.23	1.1	.27	−1.37	1.1	.21
Female	−0.57	0.85	.51	−0.65	0.86	.45	−0.54	0.86	.53
SWD	−1	0.89	.26	−0.92	0.9	.3	−0.93	0.9	.3
Grade level	0.79	1.04	.45	0.53	1.04	.61	0.71	1.03	.49
Language	−2.52	1.26	.05	−2.25	1.27	.08	−2.47	1.25	.05
SKILL × Language	0.71	1.71	.68	0.43	1.73	.8	0.67	1.71	.7
Moderator				0.48	0.37	.2	0.18	0.38	.65
SKILL × Moderator				−0.73	0.52	.16	−0.55	0.56	.33
Language × Moderator				0.06	0.56	.92	0.62	0.56	.27
SKILL × Language × Moderator				0.32	0.76	.68	−0.6	0.85	.48
Random effects	**Var.**	**ICC**		**Var.**	**ICC**		**Var.**	**ICC**	
GMRT									
Student level	88.11	.78		86.09	.78		88.71	.78	
Teacher level	25.09	.22		24.46	.22		24.31	.22	
TOSREC									
Student level	49.19	.94		50.03	.96		50.29	.97	
School level	3.07	.06		2.17	.04		1.65	.03	

Note. "Moderator" refers to TNL Comprehension in Model 2 and TNL Production in Model 3. Est. = Estimate; *SE* = standard error; GMRT = Gates–MacGinitie Reading Test; SKILL = Supporting Knowledge in Language and Literacy; SWD = student with disability; TOSREC = Test of Silent Reading Efficiency and Comprehension; Var. = variance; ICC = interclass correlation coefficient; TNL = Test of Narrative Language.

and comprehension did not differ for bilingual and monolingual students. The effect size in the monolingual group was −0.15 (95% CI [−0.44, 0.15]). In the bilingual group, the SKILL effect was −0.08 (95% CI [−0.43, 0.28]).

Research Question 2: To What Extent Does English Narrative Language Performance at Pretest Predict Bilingual and Monolingual Students' Response to Instruction?

To test whether students' English narrative language performance at pretest predicted bilingual and monolingual students' differential response to instruction, we included three-way interactions of language status, condition, and narrative language performance at pretest (as measured by TNL Comprehension and TNL Production). The interaction terms (Models 2 and 3 in Tables 3 and 4) did not differ from 0 for any of the outcomes, suggesting that response to the SKILL intervention did not differ for bilingual and monolingual students with similar initial levels of narrative language.

Discussion

A significant educational challenge has been identifying evidence-based approaches to improving oral

Table 5. Effect size estimates for outcomes.

Outcome variable	Hedges's *g* [95% CI]
Language outcomes	
TNL Comprehension	
Monolinguals	0.00 [–0.29, 0.29]
Bilinguals	0.50 [0.13, 0.86]
TNL Production	
Monolinguals	0.44 [0.14, 0.74]
Bilinguals	0.24 [–0.11, 0.60]
MISL Written Narrative	
Monolinguals	0.48 [0.18, 0.77]
Bilinguals	0.15 [–0.20, 0.51]
MISL Oral Narrative	
Monolinguals	0.67 [0.37, 0.97]
Bilinguals	0.53 [0.18, 0.89]
Reading outcomes	
GMRT	
Monolinguals	–0.02 [–0.31, 0.28]
Bilinguals	–0.03 [–0.33, 0.38]
TOSREC	
Monolinguals	–0.15 [–0.44, 0.15]
Bilinguals	–0.08 [–0.43, 0.28]

Note. CI = confidence interval; TNL = Test of Narrative Language; MISL = Monitoring Indicators of Scholarly Language; GMRT = Gates–MacGinitie Reading Test; TOSREC = Test of Silent Reading Efficiency and Comprehension.

language outcomes for bilingual students with developing English language proficiency in order to reduce the persistent achievement gap between bilingual students and their monolingual peers (National Academies of Sciences, Engineering, and Medicine, 2017). Narrative language ability is a particularly high target priority for bilinguals to enhance their language and literacy development (Uccelli & Páez, 2007). Considering the fundamental role of oral language for literacy development and the need among SLPs for effective interventions for improving language, we conducted a secondary analysis from an RCT examining the effects of a narrative language program (S. L. Gillam et al., 2023). The overarching RCT by S. L. Gillam et al. (2023) revealed that students in the SKILL condition significantly outperformed students in the BAU condition on standardized and experimental measures of oral and written narrative language development. The study reported here addressed research questions about the relative effects of SKILL for bilingual and monolingual students on English narrative language outcomes and the extent to which students' narrative language performance at pretest predicted their response to intervention. These findings hold important implications for SLPs given that evidence-based Tier 2 approaches to identification of language impairments and instruction for improving language outcomes have the potential to reduce the number of bilingual students who will require more intensive interventions.

The first research question addressed in this study was whether there were differential effects of SKILL based on students' language status. The moderation results demonstrated that the effect of SKILL did not significantly differ for bilingual and monolingual students on three of the four measures of narration. Specifically, we did not find significant differences between bilingual and monolingual students on two measures of narrative production and written narratives, although the pattern of effect sizes favored monolingual students. There was one statistically significant difference between bilingual and monolingual students. On the Narrative Comprehension subtest of the TNL, there was a significant difference in the effect of SKILL for bilingual and monolingual students, with the intervention's effect being greater for bilingual students (g = 0.51) than for monolingual students (g = 0.00).

We are interested in why the effects of SKILL may be different for bilingual and monolingual students on a measure of narrative comprehension. We are not aware of previous research that examined the differential response of school-age bilingual and monolingual students to a narrative language intervention. Among preschool children, Spencer et al. (2020) reported findings in their study of a Spanish–English version of another narrative language intervention (Story Champs). The authors reported that students demonstrated statistically significant improvements in proximal measures and that these effects were not different based on student primary language status. This finding aligns with the pattern of findings from our study—that the positive effects of a narrative language intervention are similar for EBs and monolinguals. Absent further research that replicates the finding that EBs make greater gains in narrative comprehension compared to monolinguals, we are hesitant to speculate why EBs made greater gains in narrative comprehension compared to monolingual students. Future research, specifically studies that contrast bilingual and monolingual students, may consider collecting data on a variety of linguistic and cognitive measure constructs that underlie language comprehension.

Nevertheless, the main effects of SKILL (S. L. Gillam et al., 2023) and the moderation results presented in this article underscore the promise of SKILL as an evidence-based approach for all students, including those who were bilingual. Findings also support the emerging body of research (e.g., R. D. Miller et al., 2018; Spencer et al., 2020) that shows that narrative language interventions lead to improved outcomes for bilingual students with developing English language proficiency on formal and informal measures of English narration. Although SKILL impacted students in the treatment condition on narrative comprehension performance, these findings did

not transfer to significant differences in reading comprehension development immediately for all students. This corresponds with previous research reviews that report interventions often yield small effect sizes when tested in rigorous designs using standardized measures of reading comprehension (Scammacca et al., 2015; Wanzek et al., 2010, 2016). It also aligns with previous research examining the effects of language-focused approaches to improving reading comprehension. In LARRC et al.'s (2019) large-scale RCT examining the effects of a language-focused intervention (which included narrative text structure instruction), the authors reported large and statistically significant effects on curriculum-based measures of vocabulary and comprehension monitoring, but not on more distal measures of reading comprehension.

There are several possible explanations for our finding. First, neither the GMRT nor the TOSREC primarily uses narrative literacy texts to assess reading comprehension. Had the reading comprehension measures more closely aligned with the narrative language genre taught within the SKILL treatment, participating students' improved narrative language proficiency may have generalized to reading comprehension. Second, the SKILL program is primarily a narrative language development program with a secondary emphasis on transfer to literacy. All of the SKILL lessons were oral; none required the participants to read. It may be that with extended opportunities for transfer from oral narrative language to text reading of narrative language, students would demonstrate improved reading comprehension. In a study comparing the effects of oral language instruction, text reading instruction, and typical instruction on the reading comprehension of students with poor comprehension, Clarke et al. (2010) determined that although both the oral language and text reading treatments were associated with significant reading comprehension outcomes in students, only the oral language instructional approach improved reading comprehension outcomes from posttest to follow-up testing. It is important to note that although the oral language treatment had a significant and sustained impact on reading comprehension, the target students demonstrated only reading comprehension difficulties (not word reading problems) and were monolingual.

A third explanation for why SKILL led to improvements in narrative language but not in reading comprehension may be due to the presence of word reading difficulties in this sample. To better understand our sample, we administered the TOWRE-2 SWE subtest to all participating students at pretest. Results showed that although students were screened on the basis of underperformance on narrative language and reading comprehension measures, both bilingual (standard score: M = 85.89, SD = 14.40) and monolingual (composite score: M = 84.02, SD = 15.13) students demonstrated difficulties in word reading. Given the importance of word reading and its impact on reading comprehension, it is possible that the effects of SKILL on reading comprehension were impeded due to students' code-based difficulties. For these students to realize the benefits of SKILL on a reading comprehension outcome measure, SKILL may need to be coupled with word reading instruction.

Our second research question addressed whether the effects of SKILL would vary based on bilingual students' initial English language performance. The findings revealed that initial English language performance did not moderate the effects of treatment. This is an important finding because it suggests that bilingual students do not need a certain threshold of English language proficiency to benefit from the SKILL intervention. We consider this finding compelling as there is a broad research base to suggest that initial performance on English language and reading-related tasks is a strong predictor of response to intervention (Stuebing et al., 2015). Previous research specifically related to bilingual students revealed that students' listening comprehension and word reading at the beginning of the school year significantly impacted their response to intervention (Vaughn et al., 2019). For instance, results from Vaughn et al. (2019) showed a significant three-way interaction such that (a) bilingual students outperformed their monolingual peers when students had very low word reading at pretest and (b) monolinguals outperformed bilinguals when students had relatively high word reading. The authors recommended that preintervention skill profiles may need to be interpreted differently for bilingual and monolingual students with serious reading difficulties. The fact that pretest English language performance did not have a significant effect on treatment outcomes suggests that monolingual and bilingual students with a wide range of language abilities can profit from instruction with the SKILL program.

Limitations

We acknowledge that there are limitations to this research. For one, the bilingual students in our study represented a diverse group of students with varying degrees of primary language (predominantly Spanish) and English language proficiency. Due to time constraints, we were only able to administer a limited battery of assessments in English, which focused primarily on narration. We relied on parent reports of students' home language use and proficiency. Although this is reflective of how school personnel determine whether a child is considered an English learner, this limits our understanding of the home language proficiency of EBs in this study. We acknowledge that there may have been some children who speak

another language at home (not their primary language) but were categorized as monolingual. This underscores the need to learn even more about the home language environment. We also recognize that our measures of narration did not address the full constellation of English language knowledge that comprises one's English language proficiency. S. L. Gillam et al. (2013) reported evidence to support the validity of the TNL for assessing narrative abilities in bilingual children receiving instruction in English for a year or more. However, not all of the measures have been fully studied with EBs, so there may be measurement error in the testing of EBs. Moreover, it is possible that the results may have differed had we used additional measures of initial English language proficiency. Finally, we were unable to examine whether the impact of SKILL was moderated by home language proficiency and instead used pretest performance on the TNL as an indicator of initial language proficiency. Future work should consider using multiple measures of English language proficiency and home language proficiency at pretest to further examine whether SKILL impacts bilingual students with varying degrees of first- and second-language proficiency. Another potential limitation of this study relates to generalizability. In this RCT, members of the research staff provided SKILL instruction, although 80% of the interventionists were former school-based SLPs. Further research is required to understand the extent to which these findings replicate when implementers are not members of the research staff.

Practical Implications

This study provides support for the use of SKILL, a Tier 2 narrative language–based approach to improving English narrative language and comprehension for bilingual students. There are several practical implications from this study that are worthy of consideration. First, it may be valuable for SLPs to consider using systematically developed and explicitly taught approaches such as SKILL that use a literacy-based approach to teach critical components of narrative macrostructure and microstructure in order to promote the narrative language development of monolingual and bilingual students. Second, this study provides evidence that across most narrative outcomes, SKILL is similarly effective for monolingual and bilingual students, which suggests that bilingual students do not need a unique educational program to make gains in this area (although it is worth noting that SKILL was intentionally designed with consideration for the needs of EBs). Although our study does not preclude the possibility that dual language approaches to intervention may be equally effective as or even more effective than English-only approaches, this study does suggest that an English-only intervention leads to improved outcomes for bilingual students with developing English language proficiency. This has important implications for practice given the particular shortage of bilingual SLPs (ASHA, 2018).

As with all efficacy trials, implications are most relevant to those who resemble the participant sample. In our study, the findings pertain to EBs with at least 1 year of English language instruction, most of whom spoke Spanish at home. Performance data at pretest on the TNL and parent report information suggest that children varied in their English and home language ability. One may question whether all of the EBs in this study were genuinely at risk for language and literacy difficulties or if some of their performance on the English screening measures (TNL and GMRT) reflected their limited exposure to English. Academic difficulties arising from limited exposure to a second language and those stemming from underlying neurobiological issues that underlie disabilities present similarly and are difficult to parse using study screening procedures. They are also difficult to dissect in practice. We believe it is preferable to offer a brief Tier 2 intervention to EBs who might not necessarily require it rather than to withhold a language-based intervention for a child who would have derived benefit from it. In other words, we consider Type II error (denying a child of potential benefits with the hope that their English language skills will eventually catch on) more problematic than Type I error (providing intervention to a child who may not have needed it). Offering an evidence-based Tier 2 intervention may in fact be the most effective approach for disentangling issues related to language exposure and true language and literacy difficulties.

Third, a recent observation study revealed that elementary teachers spend relatively little instructional time on narrative language instruction despite its importance in the curriculum as the focus of progressive state standards (Hall et al., 2021). Given the results of this study, classroom teachers may consider augmenting the decoding and reading comprehension instructional practices with narrative language lessons associated with improvement for monolingual and bilingual students. Universal practices may play a critical role in prevention. In an early-stage, nonrandomized study of the SKILL program, students who received whole-class narrative instruction outperformed children in a comparison classroom that followed the same curriculum with the exception of the narrative instruction on measures of narrative and vocabulary skills (S. L. Gillam et al., 2014).

Finally, our findings indicate that the SKILL program may be particularly useful as an intervention within a multitiered system. The SKILL program is an evidence-based Tier 2 practice that can be used with students with

a wide range of language and reading abilities. SLPs, classroom teachers, and special educators can expect students to make one half of a standard deviation of improvement (approximately 7 standard score points) on the TNL-2 following thirty-two 30-min small-group lessons. Students who make such improvements may not be good candidates for special education services even though they appear to be at-risk on other English language and literacy assessments. However, those students who fail to improve on measures of oral narration after the SKILL intervention may be good candidates for Tier 3 instruction.

Future Research

Results of this study demonstrate promise for the implementation of SKILL for improving the narrative language development of all learners with statistically significant and practically meaningful effects for students who are bilingual. It may be valuable to further enhance the link between the narrative language focus of SKILL and narrative literacy development. The simple view of reading (Gough & Tunmer, 1986) posits that enhancements on broader language processes would be associated with improvements in reading comprehension. However, we did not observe this transfer from improved narrative language to reading comprehension. We propose that future research might further enhance this link between narrative oral language and narrative literacy comprehension and determine the effects on both language and literacy development for students with English language and literacy difficulties. Future studies of the validity of SKILL as a tool in RTI decision making need to be conducted if SKILL is to be used in an response to intervention (RTI) context. Specifically, studies should determine the best cut scores for determining those children who need Tier 3 interventions. Finally, although our results suggest that SKILL, as currently constructed and tested, led to improved outcomes for bilingual students, future SKILL research should seek to validate RTI decisions from what bilingual students may consider, examining the effects of augmenting the SKILL curriculum to enhance and leverage students' primary language knowledge. Previous research grounded in theories of cross-linguistic transfer (e.g., Cummins, 1991) consistently shows that bilingual students' second-language acquisition is affected by their primary language knowledge (e.g., Leider et al., 2013; Proctor et al., 2006) and that bilingual students with developed primary language knowledge are more likely to progress in their second language (Genesee & Geva, 2006). Enriching the SKILL instructional lessons to ensure teachers facilitate cross-linguistic transfer and leverage bilingual funds of linguistic and cultural knowledge may enhance outcomes for bilingual students.

Data Availability Statement

The de-identified dataset generated for the current study are available from the corresponding author on reasonable request.

Acknowledgments

This research was supported by Institute of Education Sciences Grant R305A170111, awarded to Ronald B. Gillam. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences or the U.S. Department of Education.

References

Akobeng, A. K. (2005). Understanding randomised controlled trials. *Archives of Disease in Childhood, 90*(8), 840–844. https://doi.org/10.1136/adc.2004.058222

American Speech-Language-Hearing Association. (2018). *2018 Schools survey. Survey summary report: Numbers and types of responses, SLPs*. http://www.asha.org

Armon-Lotem, S., Restrepo, M. A., Lipner, M., Ahituv-Shlomo, P., & Altman, C. (2021). Vocabulary gains in bilingual narrative intervention. *Language, Speech, and Hearing Services in Schools, 52*(1), 436–448. https://doi.org/10.1044/2020_LSHSS-19-00117

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50–57. https://doi.org/10.1111/j.1540-5826.2005.00120.x

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Erlbaum.

Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., Gersten, R., Haymond, K., Kieffer, M. J., Linan-Thompson, S., & Newman-Gonchar, R. (2014). *Teaching academic content and literacy to English learners in elementary and middle school (NCEE 2014-4012)*. National Center for Education Evaluation and Regional Assistance. http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=19

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bedore, L. M., Pena, E. D., Fiestas, C., & Lugo-Neris, M. J. (2020). Language and Literacy Together: Supporting grammatical development in dual language learners with risk for language and learning difficulties. *Language, Speech, and Hearing Services in Schools, 51*(2), 282–297. https://doi.org/10.1044/2020_LSHSS-19-00055

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300.

Berman, R. A., & Slobin, D. I. (Eds.). (2013). *Relating events in narrative: A crosslinguistic developmental study*. Psychology Press. https://doi.org/10.4324/9780203773512

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwantoro, S. (2005). *The new demography of America's*

schools: Immigration and the No Child Left Behind Act. Urban Institute.

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology, 104*(1), 166–181. https://doi.org/10.1037/a0025323

Champion, T. B., Hyter, Y. D., McCabe, A., & Bland-Stewart, L. M. (2003). "A matter of vocabulary": Performances of low-income African American Head Start children on the Peabody Picture Vocabulary Test–III. *Communication Disorders Quarterly, 24*(3), 121–127. https://doi.org/10.1177/15257401030240030301

Cho, E., Capin, P., Roberts, G., Roberts, G. J., & Vaughn, S. (2019). Examining sources and mechanisms of reading comprehension difficulties: Comparing English learners and non-English learners within the simple view of reading. *Journal of Educational Psychology, 111*(6), 982–1000. https://doi.org/10.1037/edu0000332

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*(8), 1106–1116. https://doi.org/10.1177/0956797610375449

Cleave, P. L., Girolametto, L. E., Chen, X., & Johnson, C. J. (2010). Narrative abilities in monolingual and dual language learning children with specific language impairment. *Journal of Communication Disorders, 43*(6), 511–522. https://doi.org/10.1016/j.jcomdis.2010.05.005

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes, 51*(1–2), 26–49. https://doi.org/10.1080/0163853X.2013.855107

Cosentino de Cohen, C., Deterding, N., & Clewell, B. C. (2005). *Who's left behind? Immigrant children in high and low LEP schools*. Urban Institute.

Coyne, M. D., Little, M., Rawlinson, D. A., Simmons, D., Kwok, O. M., Kim, M., Simmons, L., Hagan-Burke, S., & Civetelli, C. (2013). Replicating the impact of a supplemental beginning reading intervention: The role of instructional context. *Journal of Research on Educational Effectiveness, 6*(1), 1–23. https://doi.org/10.1080/19345747.2012.706694

Cummins, J. (1991). Language development and academic learning. In L. M. Malave & G. Duquette (Eds.), *Language, culture, and cognition* (pp. 161–175). Multilingual Matters.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods, 12*(2), 121.

Espinosa, L. (2006). The social, cultural, and linguistic components of school readiness in young Latino children. In L. M. Beaulieu (Ed.), *The social–emotional development of young children from diverse backgrounds*. National Black Child Development Institute Press.

Farquharson, K., Therrien, M., Barton-Hulsey, A., & Brandt, A. F. (2021). How to recruit, support, and retain speech-language pathologists in public schools. *Journal of School Leadership, 32*(3), 225–245.

Favot, K., Carter, M., & Stephenson, J. (2021). The effects of oral narrative intervention on the narratives of children with language disorder: A systematic literature review. *Journal of Developmental and Physical Disabilities, 33*(4), 489–536. https://doi.org/10.1007/s10882-020-09763-9

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., Furgeson, J., Hayes, L., Henke, J., Justice, L., Keating, B., Lewis, W., Sattar, S., Streke, A., Wagner, R., & Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade (NCEE 2016-4008)*. National Center for Education Evaluation and Regional Assistance.

Fry, R., & Gonzales, F. (2008). *One-in-five and growing fast: A profile of Hispanic public school students*. Pew Hispanic Center.

Gándara, P. (2015). Is there really a labor market advantage to being bilingual in the U.S.? *ETS Research Report Series, 2015*(2), 1–34. https://doi.org/10.1002/ets2.12054

Garrisi, K., King, C. J., Mullin, L. J., & Gaab, N. (2020). *General Recommendations and Guidelines for Remote Assessment of Toddlers and Children, in Response to the COVID-19 Pandemic,* Open Science Framework Preprints. https://doi.org/10.31219/osf.io/wg4ef

Genesee, F., & Geva, E. (2006). Cross-linguistic relationships in working memory, phonological processes, and oral language. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 175–183). Erlbaum.

Gillam, R. B., & Pearson, N. A. (2017). *Test of Narrative Language–Second Edition*. Pro-Ed.

Gillam, R. B., Peña, E. D., Bedore, L. M., Bohman, T. M., & Mendez-Perez, A. (2013). Identification of specific language impairment in bilingual children: I. Assessment in English. *Journal of Speech, Language, and Hearing Research, 56*(6), 1813–1823. https://doi.org/10.1044/1092-4388(2013/12-0056)

Gillam, S., Hartzheim, D., Studenka, B., Simonsmeier, V., & Gillam, R. (2015). Narrative intervention for children with autism spectrum disorder (ASD). *Journal of Speech, Language, and Hearing Research, 58*(3), 920–933. https://doi.org/10.1044/2015_JSLHR-L-14-0295

Gillam, S. L., Fargo, J., Petersen, D. B., & Clark, M. (2012). Assessment of structure dependent narrative features in modeled contexts: African American and European American children. *English Linguistics Research, 1*(1), 1–17. https://doi.org/10.5430/elr.v1n1p1

Gillam, S. L., & Gillam, R. B. (2016). Narrative discourse intervention for school-aged children with language impairment. *Topics in Language Disorders, 36*(1), 20–34.

Gillam, S. L., Gillam, R. B., Fargo, J. D., Olszewski, A., & Segura, H. (2017). Monitoring Indicators of Scholarly Language: A progress-monitoring instrument for measuring narrative discourse skills. *Communication Disorders Quarterly, 38*(2), 96–106. https://doi.org/10.1177/1525740116651442

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

Gillam, S. L., Olszewski, A., Squires, K., Wolfe, K., Slocum, T., & Gillam, R. B. (2018). Improving narrative production in children with language disorders: An early-stage efficacy study of a narrative intervention program. *Language, Speech, and Hearing Services in Schools, 49*(2), 197–212. https://doi.org/10.1044/2017_LSHSS-17-0047

Gillam, S. L., Vaughn, S., Roberts, G., Capin, P., Fall, A.-M., Israelsen-Augenstein, M., Holbrook, S., Wada, R., Hancock, A., Fox, C., Dille, J., Magimairaj, B. M., & Gillam, R. B. (2023). Improving oral and written narration and reading comprehension of children at-risk for language and literacy difficulties: Results of a randomized clinical trial. *Journal of Educational Psychology, 115*(1), 99–117.

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

Greenhalgh, K. S., & Strong, C. J. (2001). Literate language features in spoken narratives of children with typical language and children with language impairments. *Language, Speech, and Hearing Services in Schools, 32*(2), 114–125. https://doi.org/10.1044/0161-1461(2001/010)

Hall, C., Capin, P., Vaughn, S., Gillam, S. L., Wada, R., Fall, A. M., Roberts, G., Dille, J. T., & Gillam, R. B. (2021). Narrative instruction in elementary classrooms: An observation study. *The Elementary School Journal, 121*(3), 454–483. https://doi.org/10.1086/712416

Heilmann, J. J., Rojas, R., Iglesias, A., & Miller, J. F. (2016). Clinical impact of wordless picture storybooks on bilingual narrative language production: A comparison of the 'Frog' stories. *International Journal of Language & Communication Disorders, 51*(3), 339–345. https://doi.org/10.1111/1460-6984.12201

Hernandez, D. J., Denton, N. A., & Macartney, S. E. (2008). Children in immigrant families: Looking to America's future. *Social Policy Report, 22*(3), 1–24. https://doi.org/10.1002/j.2379-3988.2008.tb00056.x

Iluz-Cohen, P., & Walters, J. (2012). Telling stories in two languages: Narratives of bilingual preschool children with typical and impaired language. *Bilingualism: Language and Cognition, 15*(1), 58–74. https://doi.org/10.1017/S1366728911000538

Jensen, T. A. (2009). *Narrative language intervention for English language learners* [Thesis, Utah State University]. Undergraduate Honors Capstone Projects. https://digitalcommons.usu.edu/honors/11

Kenny, D. A. (2018). *Moderator variables: Introduction*. https://davidakenny.net/cm/moderation.htm

Kintsch, W. (2018). Revisiting the construction—integration model of text comprehension and its Implications for Instruction. In D. E. Alvermann, N. J. Unrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy,* (pp. 178–203). Routledge.

Kong, N. Y., & Hurless, N. (2023). Vocabulary interventions for young emergent bilingual children: A systematic review of experimental and quasi-experimental studies. *Topics in Early Childhood Special Education, 43*(1), 17–29. https://doi.org/10.1177/02711214211027625

Language and Reading Research Consortium, Jiang, H., & Logan, J. (2019). Improving reading comprehension in the primary grades: Mediated effects of a language-focused classroom intervention. *Journal of Speech, Language, and Hearing Research, 62*(8), 2812–2828. https://doi.org/10.1044/2019_JSLHR-L-19-0015

Leider, C. M., Proctor, C. P., Silverman, R. D., & Harring, J. R. (2013). Examining the role of vocabulary depth, cross-linguistic transfer, and types of reading measures on the reading comprehension of Latino bilinguals in elementary school. *Reading and Writing, 26*(9), 1459–1485. https://doi.org/10.1007/s11145-013-9427-6

Lenth, R. 2020. *emmeans: Estimated marginal means, aka least-squares means*. https://cran.r-project.org/package=emmeans

Llosa, L., Lee, O., Jiang, F., Haas, A., O'Connor, C., Van Booven, C. D., & Kieffer, M. J. (2016). Impact of a large-scale science intervention focused on English language learners. *American Educational Research Journal, 53*(2), 395–424. https://doi.org/10.3102/0002831216637348

Lovett, M. W., De Palma, M., Frijters, J., Steinbach, K., Temple, M., Benson, N., & Lacerenza, L. (2008). Interventions for reading difficulties: A comparison of response to intervention by ELL and EFL struggling readers. *Journal of Learning Disabilities, 41*(4), 333–352. https://doi.org/10.1177/0022219408317859

Lucero, A. (2018). Oral narrative retelling among emergent bilinguals in a dual language immersion program. *International Journal of Bilingual Education and Bilingualism, 21*(2), 248–264. https://doi.org/10.1080/13670050.2016.1165181

MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2001). *Gates-MacGinitie Reading Test–Fourth Edition (GMRT-4)*. Riverside Publishing.

Magimairaj, B. M., Capin, P., Gillam, S. L., Vaughn, S., Roberts, G., Fall, A. M., & Gillam, R. B. (2022). Online administration of the Test of Narrative Language–Second Edition: Psychometrics and considerations for remote assessment. *Language, Speech, and Hearing Services in Schools, 53*(2), 404–416. https://doi.org/10.1044/2021_LSHSS-21-00129

Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development, 82*(5), 1544–1560. https://doi.org/10.1111/j.1467-8624.2011.01633.x

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9*(1), 111–151. https://doi.org/10.1016/0010-0285(77)90006-8

Mayer, M. (2003). *Frog, where are you?* Penguin.

McCabe, A., & Bliss, L. S. (2005). Narratives from Spanish-speaking children with impaired and typical language development. *Imagination, Cognition and Personality, 24*(4), 331–346. https://doi.org/10.2190/CJQ8-8C9G-05LG-0C2M

McLeod, S., Verdon, S., Baker, E., Ball, M. J., Ballard, E., David, A. B., Bernhardt, B. M., Bérubé, D., Blumenthal, M., Bowen, C., Brosseau-Lapré, F., Bunta, F., Crowe, K., Cruz-Ferreira, M., Davis, B., Fox-Boyer, A., Gildersleeve-Neumann, C., Grech, H., Goldstein, B., . . . The International Expert Panel on Multilingual Children's Speech. (2017). Tutorial: Speech assessment for multilingual children who do not speak the same language(s) as the speech-language pathologist. *American Journal of Speech-Language Pathology, 26*(3), 691–708. https://doi.org/10.1044/2017_AJSLP-15-0161

Méndez, L. I., Crais, E. R., Castro, D. C., & Kainz, K. (2015). A culturally and linguistically responsive vocabulary approach for young Latino dual language learners. *Journal of Speech, Language, and Hearing Research, 58*(1), 93–106. https://doi.org/10.1044/2014_JSLHR-L-12-0221

Miller, J. F., Heilmann, J., Nockerts, A., Iglesias, A., Fabiano, L., & Francis, D. J. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research & Practice, 21*(1), 30–43. https://doi.org/10.1111/j.1540-5826.2006.00205.x

Miller, R. D., Correa, V. I., & Katsiyannis, A. (2018). Effects of a story grammar intervention with repeated retells for English learners with language impairments. *Communication Disorders Quarterly, 40*(1), 15–27. https://doi.org/10.1177/1525740117751897

Morgan, P. L., Farkas, G., Hillemeier, M. M., Mattison, R., Maczuga, S., Li, H., & Cook, M. (2015). Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. *Educational Researcher, 44*(5), 278–292.

National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting the educational success of children and youth learning English: Promising futures*. The National Academies Press.

National Assessment of Educational Progress. (2022). *The Nation's Report Card: Reading*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. http://nces.ed.gov/

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards English Language Arts*.

Nicolopoulou, A., & Trapp, S. (2018). Narrative interventions for children with language disorders: A review of practices and findings. In A. Bar-On & D. Ravid (Eds.), *Handbook of*

communication disorders: Theoretical, empirical, and applied linguistic perspectives (pp. 357–386). De Gruyter Mouton. https://doi.org/10.1515/9781614514909-018

Oller, D. K. & Pearson, B. Z. (2002). Assessing the effects of bilingualism: A background. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (Vol. 2, pp. 3–21). Multilingual Matters. https://doi.org/10.21832/9781853595721-002

Pearson, B. Z. (2002). Narrative competence among monolingual and bilingual school children in Miami. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 135–174). Multilingual Matters. https://doi.org/10.21832/9781853595721-008

Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *Journal of Speech, Language, and Hearing Research, 57*(6), 2208–2220. https://doi.org/10.1044/2014_JSLHR-L-13-0151

Pesco, D., & Gagné, A. (2017). Scaffolding narrative skills: A meta-analysis of instruction in early childhood settings. *Early Education and Development, 28*(7), 773–793. https://doi.org/10.1080/10409289.2015.1060800

Petersen, D. B. (2011). A systematic review of narrative-based language intervention with children who have language impairment. *Communication Disorders Quarterly, 32*(4), 207–220. https://doi.org/10.1177/1525740109353937

Petersen, D. B., Staskowski, M., Spencer, T. D., Foster, M. E., & Brough, M. P. (2022). The effects of a multitiered system of language support on kindergarten oral and written language: A large-scale randomized controlled trial. *Language, Speech, and Hearing Services in Schools, 53*(1), 44–68. https://doi.org/10.1044/2021_LSHSS-20-00162

Pico, D. L., Hessling Prahl, A., Biel, C. H., Peterson, A. K., Biel, E. J., Woods, C., & Contesse, V. A. (2021). Interventions designed to improve narrative language in school-age children: A systematic review with meta-analyses. *Language, Speech, and Hearing Services in Schools, 52*(4), 1109–1126. https://doi.org/10.1044/2021_LSHSS-20-00160

Price, J. R., Roberts, J. E., & Jackson, S. C. (2006). Structural development of the fictional narratives of African American preschoolers. *Language, Speech, and Hearing Services in Schools, 37*(3), 178–190. https://doi.org/10.1044/0161-1461(2006/020)

Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology, 98*(1), 159–169. https://doi.org/10.1037/0022-0663.98.1.159

Restrepo, M. A., Morgan, G. P., & Thompson, M. S. (2013). The efficacy of a vocabulary intervention for dual-language learners with language impairment. *Journal of Speech, Language, and Hearing Research., 56*(2), 748–765. https://doi.org/10.1044/1092-4388(2012/11-0173)

Rogde, K., Hagen, Å. M., Melby-Lervåg, M., & Lervåg, A. (2019). The effect of linguistic comprehension instruction on generalized language and reading comprehension skills: A systematic review. *Campbell Systematic Reviews, 15*(4), Article e1059. https://doi.org/10.1002/cl2.1059

Rothstein, R. (2017). *The color of law: A forgotten history of how our government segregated America*. Liveright Publishing.

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

Silliman, E. R., Bahr, R. H., Brea, M. R., Hnath-Chisolm, T., & Mahecha, N. R. (2002). Spanish and English proficiency in the linguistic encoding of mental states in narrative retellings. *Linguistics and Education, 13*(2), 199–234. https://doi.org/10.1016/S0898-5898(01)00062-6

Simon-Cereijido, G., & Gutiérrez-Clellen, V. F. (2009). A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains. *Applied Psycholinguistics, 30*(2), 315–337. https://doi.org/10.1017/S0142716409090134

Spencer, T. D., Moran, M., Thompson, M. S., Petersen, D. B., & Restrepo, M. A. (2020). Early efficacy of multitiered dual-language instruction: Promoting preschoolers' Spanish and English oral language. *AERA Open, 6*(1). https://doi.org/10.1177/2332858419897886

Spencer, T. D., & Petersen, D. B. (2020). Narrative intervention: Principles to practice. *Language, Speech, and Hearing Services in Schools, 51*(4), 1081–1096. https://doi.org/10.1044/2020_LSHSS-20-00015

Spencer, T. D., Petersen, D. B., & Adams, J. L. (2015). Tier 2 language intervention for diverse preschoolers: An early-stage randomized control group study following an analysis of response to intervention. *American Journal of Speech-Language Pathology, 24*(4), 619–636. https://doi.org/10.1044/2015_AJSLP-14-0101

Squires, K. E., Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., Bohman, T. M., & Gillam, R. B. (2014). Story retelling by bilingual children with language impairments and typically developing controls. *International Journal of Language & Communication Disorders, 49*(1), 60–74. https://doi.org/10.1111/1460-6984.12044

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. *New Directions in Discourse Processing, 2*(1979), 53–120.

Stuebing, K. K., Barth, A. E., Trahan, L. H., Reddy, R. R., Miciak, J., & Fletcher, J. M. (2015). Are child cognitive characteristics strong predictors of responses to intervention? A meta-analysis. *Review of Educational Research, 85*(3), 395–429. https://doi.org/10.3102/0034654314555996

Thordardottir, E. (2010). Towards evidence-based practice in language intervention for bilingual children. *Journal of Communication Disorders, 43*(6), 523–537. https://doi.org/10.1016/j.jcomdis.2010.06.001

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency–Second Edition (TOWRE-2)*. Pearson Clinical Assessment.

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly, 50*(3), 337–356. https://doi.org/10.1002/rrq.104

Uccelli, P., & Páez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools, 38*(3), 225–236. https://doi.org/10.1044/0161-1461(2007/024)

Ukrainetz, T. A. (2006). Teaching narrative structure: Coherence, cohesion, and captivation. In T. A. Ukrainetz (Ed.), *Contextualized language intervention: Scaffolding PreK-12 literacy achievement* (pp. 195–246). Pro-Ed.

Vaughn, S., Martinez, L. R., Wanzek, J., Roberts, G., Swanson, E., & Fall, A. M. (2017). Improving content knowledge and comprehension for English language learners: Findings from a randomized control trial. *Journal of Educational Psychology, 109*(1), 22.

Vaughn, S., Roberts, G., Capin, P., Miciak, J., Cho, E., & Fletcher, J. M. (2019). How initial word reading and language

skills affect reading comprehension outcomes for students with reading difficulties. *Exceptional Children, 85*(2), 180–196. https://doi.org/10.1177/0014402918782618

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *TOSREC: Test of Silent Reading Efficiency and Comprehension*. Pro-Ed.

Wang, R., & Ware, J. H. (2013). Detecting moderator effects using subgroup analyses. *Prevention Science, 14*(2), 111–120. https://doi.org/10.1007/s11121-011-0221-x

Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-analyses of the effects of tier 2 type reading interventions in Grades K–3. *Educational Psychology Review, 28*(3), 551–576. https://doi.org/10.1007/s10648-015-9321-7

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing, 23*(8), 889–912. https://doi.org/10.1007/s11145-009-9179-5

What Works Clearinghouse. (2020). *What Works Clearinghouse standards handbook* (Version 4.1). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/handbooks