Learning and Individual Differences

The Role of Knowledge Availability in Forming Inferences with Rural Middle Grade English Learners

--Manuscript Draft--

Manuscript Number:	LEAIND-D-20-00555R2
Article Type:	Full Length Article
Keywords:	Inference, knowledge-base, English language learners, rural, middle school
Corresponding Author:	Amy Elizabeth Barth Buena Vista University Buena Vista, UNITED STATES
First Author:	Amy Elizabeth Barth
Order of Authors:	Amy Elizabeth Barth
	Johny Daniel
	Gregory Roberts
	Sharon Vaughn
	Marcia Barnes
	Ethan Ankrum
	Heather Kincaid
Abstract:	We investigated differences in knowledge-based inferencing between rural, middle grade monolingual English-speaking students and English learners. Students were introduced to facts about an imaginary planet Gan followed by a multi-episode story about Gan. Participants were tested on the accuracy of fact recall and inferences using this knowledge at three time points (i.e., immediate, one-week, and one-month follow-up). Results show that monolingual English-speaking students significantly outperformed English learners on the inference task. Both subgroups made elaborative inferences more accurately than coherence. Students' ability to recall knowledge base facts was the strongest predictor of their ability to accurately make inferences using this knowledge at each time point.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

*Title Page WITH author details

The Role of Knowledge Availability in Forming Inferences with Rural Middle Grade English Learners

Amy E. Barth,[1] Johny Daniel,[2] Gregory Roberts, [3] Sharon Vaughn,[3] Marcia A. Barnes,[4] Ethan Ankrum,[1,] and Heather Kincaid[1]

[1]Department of Literacy, Buena Vista University

[2]Department of Special Education, Boston University

[3]Meadows Center for Preventing Educational Risk, University of Texas at Austin

[4]Vanderbilt University

Author Note

Amy E. Barth https://orcid.org/0000-0003-1392-6737

Johny Daniel https://orcid.org/0000-0002-5057-9933

Greg Roberts https://orcid.org/0000-0001-6333-7442

Sharon Vaughn https://orcid.org/0000-0001-8305-5549

Marcia Barnes https://orcid.org/0000-0002-4849-886X

Ethan Ankrum https://orcid.org/0000-0002-4810-5288

Heather Kincaid https://orcid.org/0000-0001-5242-5471

Highlights

1. English learners (ELs) require more opportunities to acquire new knowledge.
2. ELs and English-speaking students retain new knowledge at comparable levels.
3. ELs do not consistently use new knowledge to form inferences while reading.
4. Knowledge is a necessary but not sufficient condition for inferencing while reading.

REVISED Manuscript (text UNmarked) WITHOUT author details
Click here to view linked References

The Role of Knowledge Availability in Forming Inferences with Rural Middle Grade English Learners

Abstract

We investigated differences in knowledge-based inferencing between rural, middle grade monolingual English-speaking students and English learners. Students were introduced to facts about an imaginary planet Gan followed by a multi-episode story about Gan. Participants were tested on the accuracy of fact recall and inferences using this knowledge at three time points (i.e., immediate, one-week, and one-month follow-up). Results show that monolingual English-speaking students significantly outperformed English learners on the inference task. Both subgroups made elaborative inferences more accurately than coherence. Students' ability to recall knowledge base facts was the strongest predictor of their ability to accurately make inferences using this knowledge at each time point.

Keywords: Inference, knowledge-base, English language learners, rural, middle school

Introduction 1.1

The importance of comprehension increases in middle school when expectations to learn through reading grow alongside the increasingly complex language of grade-level texts. The demand to read with comprehension is even greater for English learners (ELs) who must develop proficiency in listening, speaking, reading, and writing English as a second language. ELs represent approximately 20% of US children (Ryan, 2013) and constitute the fastest growing population in US schools (Kena et al., 2015). Of concern, very few ELs read at proficient levels (6%) by eighth grade (National Center for Education Statistics, 2017). Population data shows that many ELs attend rural, agriculturally-based community school districts where they are subject to unavailable, inaccessible, or unacceptable support services (DeLeon et al., 2003). This coupled with the poor reading performance of ELs, makes understanding how to improve reading comprehension a priority for middle grade ELs (Lesaux, 2006; 2010).

Despite varying and often conflicting views on the nature and definition of reading comprehension, the ability to form knowledge-based inferences remains the cornerstone of comprehension (Ahmed et al., 2016; Kendeou, 2015). Knowledge-based inferences integrate important information in the text with the reader's general knowledge of the topic. Knowledge-based inferences are generated because texts rarely include every detail required for understanding, remembering, and learning. Therefore, readers regularly use their world knowledge to make inferences to fill in details to construct a complete situation model of the text (Kintsch, 1988).

At least two types of knowledge-based inferences are important for building a situation model. Coherence inferences helps the reader to understand the text by "gap filling." Coherence inferences add important but unstated information so the reader can infer why an event occurred.

For example, on hearing "The captain announced that the airplane required immediate maintenance. The passengers were angry." An inference about a possible flight delay is important for understanding "why" passengers were angry. Elaborative inferences help to specify a more detailed account of the situation so that the reader may infer what an event was like (Kintsch et al., 1993). For example, inferring that the sky was full of smoke, upon hearing "The California wildfires were causing significant visibility problems," contributes to a richer understanding of the situation described in the text. Elaborative inferences are not routinely constructed to maintain textual or causal coherence but are formed to embellish story content and enrich its context.

Skilled readers regularly make coherence and elaborative inferences (e.g., Albrecht & O'Brien, 1993; Bloom et al., 1990; van den Broek et al., 2005). However, these inferences can only be formed when the requisite general knowledge necessary to form the inference is available (Cain et al., 2001). General knowledge is defined as being available if it is stored in semantic memory and can be retrieved in some but not all listening or reading contexts. Past research indicates that availability of general knowledge accounts for individual and developmental differences in inference making (Barnes et al., 1996; Barnes et al., 2015; Cain et al., 2001) and that knowledge is a major contributor to inference-making and comprehension (Ahmed et al., 2016; Cromley et al., 2009). These studies suggest that general knowledge differences are a primary source of individual differences in knowledge-based inference-making among secondary grade readers (Ahmed et al., 2016).

When we examine the wider literature on inference-making, very little is known about the role of general knowledge on inference-making among English learners (ELs). Cross-sectional and longitudinal studies consistently document that semantic knowledge in English is a

significant area of difficulty for ELs (August & Shanahan, 2006; Carlo et al., 2004; Galloway & Lesaux, 2015; Goldenberg, 2011; Kieffer, 2010; Lesaux et al., 2010; Nakamoto et al., 2008; Proctor et al., 2005; Reed et al., 2016; Uccelli et al., 2015; Verhoeven, 2011) and is associated with their reading comprehension difficulties (Mancilla-Martinez & Lesaux, 2010). These studies provide clear evidence that general knowledge in English accounts for individual and developmental differences for a variety of lower-level and higher-level reading skills (Proctor et al., 2019). Further, a growing body of evidence suggests that ELs have significantly fewer first or second language word meanings stored in their lexicon relative to their monolingual English-speaking peers (e.g., Hoff et al., 2012; Mancilla-Martinez & Vagh, 2013). This is not because of an inherently lower propensity for the acquisition of new word meanings (Bedore & Peña, 2008), rather it is associated the distributed nature of language exposure across first and second languages (Dijkstra & Van Heuven, 2002; Kroll & Stewart, 1994). Yet, no study has investigated the effect of general knowledge on inference-making among ELs after the new knowledge is trained to the same level of mastery as their monolingual English speaking peers and where knowledge affects both the probability that an inference will be made as well as the strength with which an inference is encoded (Kintsch, 1994; McKoon & Ratcliff, 1992).

Therefore, the purpose of this study is to examine the role of general knowledge in coherence and elaborative knowledge-based inferencing among middle grade ELs attending under-resourced schools in a rural setting. We first synthesize the research on the role of knowledge in inferencing among monolingual English-speaking readers and then present findings from several international studies of inference-making conducted with emergent ELs from a range of cultural and linguistic backgrounds.

The Role of General Knowledge in Inference-making 1.1.1

Two methods have primarily been used to examine the role of general knowledge in inference-making. The first method adapted reading materials to individual students consistent with their knowledge of the topic. The second selected groups of students differing in their domain knowledge for selected texts. Evidence from both lines of research converge to show that larger and denser networks of general knowledge enable activation and integration of this knowledge when forming inferences (e.g., Kendeou, 2015; Kendeou & van den Broek, 2005, 2007; McNamara et al., 2007), leading to higher levels of reading comprehension and improved memory for the text's central ideas (e.g., Barnes et al.,, 2015; Chiesi et al., 1979; Haenggi & Perfetti, 1992; Kendeou & van den Broek, 2007; Miller & Keenan, 2009; Rawson & Kintsch, 2004; Recht & Leslie, 1988; Samuelstuen & Braten, 2005). Further, studies have demonstrated that revisions to text that increase their structural and explanatory coherence (e.g., that supply relevant background knowledge and explain important causal relationships) lead to improved reading comprehension among elementary grade students (Beck et al., 1991) and college students (Britton & Gulgoz, 1991). However, even when texts provide knowledge support to low knowledge readers, these readers may still experience comprehension challenges. These challenges arise because access to the broad range of semantic connections needed to integrate information within and between sentences with related knowledge stored in long-term memory are not sufficient (Barnes et al., 2015; Kintsch, 1994).

The influence of general knowledge on inference making has also been studied using an experimental procedure that manipulated the reader's knowledge base to determine if inference difficulties may be accounted for by a "general knowledge" gap. In these studies, a new knowledge base was taught to all participants and the only inferences required (i.e., coherence and elaborative inferences) drew on this newly learned knowledge (Barnes & Dennis, 1998;

Barnes et al., 1996; Cain et al., 2001). These studies compared the inference skills of younger versus older children and children with reading difficulties or disabilities versus typically developing peers. Barnes and colleagues report that younger children (Barnes et al., 1996), students with disabilities (Barnes & Dennis, 1998), and students with reading comprehension difficulties (Cain et al., 2001) required more learning trials to acquire the novel knowledge base. When the knowledge base was re-tested after reading the story, retention of the knowledge base was comparable across typically developing students in the elementary and middle grades (Barnes et al., 1996), but students with reading disabilities (Barnes & Dennis, 1998) and reading difficulties (Cain et al., 2001) recalled fewer knowledge base elements. In addition, when knowledge is equally available to students of different ages (Barnes et al., 1996) as well as skilled and less skilled comprehenders (Cain et al., 2001), developmental and individual differences in knowledge-based inferencing remain. Last, accuracy in forming knowledge-based inferences also varied by reading context (i.e., during reading versus after reading) (Barnes et al., 1996; Cain et al., 2001). Results demonstrated that the difficulties of younger students in forming knowledge-based inferences during text processing (a more complex processing situation) can be accounted for by differences in the ability to form inferences in isolation from text. Younger students presented a basic limitation in knowledge-based inference making irrespective of knowledge and processing context (Barnes et al., 1996). These studies consistently show that knowledge is associated with differences in inferencing, but knowledge availability does not ensure that inferences will be accurately made.

Inference-making and Second Language Learners 1.1.2

International studies examining inferencing skill among bilingual (BL) and second language learners (SL) are rare (see Martinez-Lincoln et al., 2021) and generally suffer from

small sample sizes and measurement issues that preclude generalization to ELs attending US schools. Recognizing these weaknesses, results suggest that typically developing bilingual students with adequate first and second language skills do not differ significantly from monolingual students in general inference-making skill (Bonifacci et al., 2017; Bonifacci & Tobia, 2016; Burgoyne et al., 2013; Geva & Massey-Garrison, 2012). Among bilingual students, SLs, and monolingual students, higher levels of reading comprehension are associated with higher levels of inferencing (Bonifacci & Tobbi, 2016; D'Angelo & Chen, 2017; Geva & Massey-Garrison, 2012 Li, 2012; Prior et al., 2014). Although inferencing among SLs is lower relative to bilingual or monolingual students (Martinez-Lincoln et al., in review), inferencing significantly improves when SLs read culturally familiar texts (Abu-Rabia 1996; August et al., 2006; Kembo, 2001) or elaborated texts that contain redundancy and definitions for low frequency words (Oh, 2001). Finally, when bilinguals read texts in their second language, relative to reading texts in their primary language, they have trouble recalling the text's central ideas (Miller & Keenan, 2011). This "centrality deficit" likely contributes to slower and less accurate inferencing while reading. In summary, knowledge is essential for inferencing across a variety of languages (Elbro & Buch-Iverson, 2013) but knowledge availability is not sufficient for adequate inferencing (Barnes et al., 1996; Cain et al., 2001)

Study Purpose 1.1.3

This study examines whether knowledge-based inference making differences exist between typically developing, monolingual English-speaking students and ELs in the middle grades attending under-resourced rural schools. We describe elaborative and coherence inference-making, where a new knowledge base is taught to all students and where the only inferences required are those that draw on this taught knowledge base. The use of a newly

learned knowledge base affords the opportunity to examine knowledge-based inference-making among ELs in the absence of a knowledge gap that may be specific to their English language development. It also levels accumulated practice of the knowledge-base across ELs and monolingual English speakers. In addition, we examined the stability of the knowledge base as well as the use of the knowledge-base to form inferences over time. Because small differences in retention or stability of knowledge can have profound effects on lexical retrieval and accessibility (Murray & Forster, 2004), we were interested in understanding if inference-making was related to the stability of the knowledge base over time and if group differences existed. Also, few studies collect follow up data to determine long term intervention effects (Daniel et al., 2020). Third, this study sought to examine the inference making skills of ELs and monolingual English-speaking students while comprehending a story versus afterwards. If ELs struggle to accurately integrate relevant information from text with the knowledge-based fact after listening to the story and when the question helps to cue the inference, then any inference difficulties identified while reading text may reflect more basic integration challenges. Finally, we were interested in examining knowledge-based inference making among *typically developing* ELs to understand if individual differences are negated once the requisite knowledge is accumulated at comparable levels of accurate recall. Findings would then suggest that differences between ELs and monolingual English speakers, which have been reported in past studies examining higher-level language skills (e.g., Proctor et al., 2019), are not due to a basic limitation in inferencing but rather are due to differences in English language knowledge.

Research questions include: (1) Are there differences in learning and retention of information from a novel knowledge base between ELs and monolingual English students immediately and over time? (2) How do coherence and elaborative inferences compare in ELs

and mono-lingual English students immediately after learning the novel knowledge base and over time? (3) Does inference making in the context of text comprehension differ from inference making directly elicited after the story concludes?

Material and Methods 1.2

Participants 1.2.1

This study was conducted in a middle school located in a rural midwestern area of the United States. All participants were typical readers, able to read grade-level text fluently. Human subjects approval was granted by the Institutional Review Board/Institutional Animal Care and Use Committee as well as accordance with human subjects guidelines and principles. All materials were in Spanish and English. Language translation assistance was also provided by the school district for families who spoke a language other than Spanish or English. All participants required consent from parents and/or legal guardians and gave both oral and written assent before participating in this study. *Typical* and *fluently* were evaluated using the Gates MacGinitie Reading Comprehension subtest (GMRT-4; MacGinitie et al., 2000) and the Fastbridge CBM Reading Fluency assessments (FAST; Christ, 2015), respectively. A total of 145 students in Grades 5 and 6 met the inclusion criteria. Approximately 60% of participants were classified as ELs, 52% as Hispanic/Latino, 26% Caucasian, and 16% Asian, with the remaining students classified as either multi-racial or non-disclosed. See Table 1 for student demographic and descriptive statistics on selection measures.

A total of 125 students in Grade 5 and 6 did not meet inclusion criteria. In Grade 5, 17 monolingual speaking students and 57 ELs were excluded from participation in the study. In Grade 6, 13 monolingual speaking students and 38 ELs were excluded. Average reading comprehension standard scores and average fluency (i.e., WCPM) scores for the excluded Grade

5 sample was 90.10 and 121.06, respectively, and for the Grade 6 sample scores were 85.87 and 127.30, respectively. No participant disclosed any physical or cognitive disabilities.

The GMRT requires that students read passages silently and answer multiple-choice questions that tap literal understanding of text, vocabulary knowledge, and inference-making. To ensure all students had adequate reading comprehension skills, the grade 5 and 6 cut-off was a standard score of 85. Internal consistency reliability ranges from .91 to .93, and alternate form reliability is reported as .80 to .87 (MacGinitie et al., 2000).

The FAST is an evidence-based, individually administered, one-minute reading fluency assessment for students in Grades 1–6. After reading a text for one minute, words correct per minute (WCPM) was calculated. To ensure all students had adequate reading fluency skills, the Grade 5 and 6 benchmarks was 133 WCPM and 142 WCPM, respectively. Internal consistency ranges from .88 to .93, the test-retest reliability ranges from .92 to .94, and alternate form reliability ranges from .78 to .95. To qualify for the study, students had to score 85 > on GMRT and score above the FAST CBM grade-level fluency benchmark.

Design and Procedures 1.2.3

This study was part of a larger study examining knowledge-based inferences among rural, ELs and monolingual English speakers in the middle grades. The materials were modified from those used by Barnes and colleagues (1996), with all stimuli presented on computer using E-Prime and Chronos software (Psychology Software Tools, 2016; Schneider et al., 2012). Administration on the computer ensured high fidelity of testing across participants. All students were tested individually by a trained undergraduate laboratory assistant in a quiet space identified by the middle school literacy coach. All students completed the same three-phase procedure (See Figure 1).

(1) Phase 1: Students learned a new knowledge base that consisted of 14 facts about an imaginary planet called Gan.

(2) Phase 2: Students listened to a seven-episode story about Gan. Following each episode, students were asked 1 coherence and 1 elaborative inference question.

(3) Phase 3: Retention of the taught knowledge base was retested and performance on the coherence and elaborative inferences was reassessed immediately and one-week and one-month later.

Phase 1: Learning the Knowledge Base 1.2.4

The knowledge base included 14 facts that ascribed a property related to the people, environment, or objects on the invented world of "Gan." Facts were presented aurally by the computer through headphones. After listening to the knowledge base, acquisition was assessed using a forced-choice picture-recognition task and a verbal recall task.

Forced-choice picture-selection task. Students completed a forced-choice picture-selection test comprised of 14 trials, one trial for each fact in the knowledge base. For each trial, students were presented four pictures on the computer that included one correct answer and three distractors (see Appendix B). For example, the test asked the student to "Point to the picture of some turtles on Gan." The student's task was to select the picture corresponding to a true property about Gan by pressing the correct button on a four-button keypad. In this example, the student would select picture "A". A score of 14 indicates that the student correctly answered all questions the first time. A score of 15 indicates that the student missed 1 item and that upon retesting answered it correctly.

Verbal recall task. The verbal recall task comprised of 14 questions was designed to assess students' memory of the knowledge base. Questions were aurally presented by the

computer through headphones, with students' oral responses recorded by the computer and scored as correct or incorrect by the laboratory assistant. For example, a knowledge base item is, "The rivers and ponds on Gan are filled with Orange Juice." The associated verbal recall question is, "What are the ponds on Gan filled with?" The correct answer is "orange juice." A score of 14 indicates that the student correctly answered all questions the first time. A score of 15 indicates that the student missed 1 item and that upon retesting answered it correctly.

In both the forced-choice picture-selection task and verbal recall task, incorrect answers were corrected immediately and retested after all items had been presented. Items recalled incorrectly were presented by the computer until mastered. No time limit was given to answer each question, but students were encouraged to work quickly and accurately.

Phase 2: Story episodes and inference questions 1.2.5

Next students listened to a seven-episode story and 21 questions that were presented aurally by the computer through headphones, one episode at a time. Episodes ranged in length from 142–169 words and included one coherence and one elaborative inference question. Students' verbal answers to the questions were recorded by the computer and scored by the laboratory assistant. When a question was answered incorrectly or incompletely, the laboratory assistant would say, "tell me more about that," to elicit a more complete answer. See Appendix C for an example story episode and associated questions. Analyses were conducted using the total number of inferences correctly formed when the question was first asked or after the prompt for further information ("tell me more about that").

Phase 3: Retention and Use of the Knowledge Base to Form Inferences After the Story 1.2.6

Retention of the knowledge base. Retention of the knowledge base was re-tested immediately after the story episode, one-week and one-month later using the verbal recall task

described above, but no feedback or re-testing to criterion occurred. Follow-up assessments of knowledge base retention measured if the knowledge base was equally available for ELs and monolingual English speakers.

Use of the knowledge base to form inferences. Coherence and elaborative inference making was also assessed at the previously mentioned three follow-up time points. The coherence and elaborative inference questions were cued in that they repeated information from the seven-episode story that the student needed to accurately form the inference (Appendix D).

Analytic Approach 1.2.7

Data for this study were taken from a larger, randomized controlled trial designed to evaluate the effect of teaching the knowledgebase under different learning conditions on the accuracy of inferences. For this study, we fit single level models for three specific time-points to estimate the effects of key covariates such as knowledge base recall accuracy, EL status, and student response to different question types. The Benjamini-Hochberg (Benjamini & Hochberg, 1995; 2000) correction was used to control for false discovery rate (FDR) associated with multiple comparisons, which include the two analyses of covariances (ANCOVAs) estimated for each of the three time points. We applied the correction separately to models that were and were not controlling for question type. The ANCOVAs control for students' pretest GMRT scores which also optimizes statistical power; the covariate was grand mean centered for ease of interpretation. Data modeling was done using R (R Core Team, 2015).

Figure 1. Models 1, 2, and 3 for the Know-it Inferential Measure.

Model 1a: $y_{immediate} = B_0 + B_{1RCpretest} + B_{2VerbalRecall\_immediate} + B_{3ELL} + \varepsilon_i$

Model 1b: $y_{immediate} = B_0 + B_{1RCpretest} + B_{2VerbalRecall\ \_immediate} + B_{3ELL} + B_{4Question\_Type} + \varepsilon_i$

Model 2a: $y_{one\text{-}week} = B_0 + B_{1RCpretest} + B_{2VerbalRecall\ \_one\text{-}week} + B_{3ELL} + \varepsilon_i$

Model 2b: $y_{one\text{-}week} = B_0 + B_{1RCpretest} + B_{2VerbalRecall\_\ one\text{-}week} + B_{3ELL} + B_{4Question\_Type} + \varepsilon_i$

Model 3a: $y_{one\text{-}month} = B_0 + B_{1RCpretest} + B_{2VerbalRecall\_one\text{-}month} + B_{3ELL} + \varepsilon_i$

Model 3b: $y_{one\text{-}month} = B_0 + B_{1RCpretest} + B_{2VerbalRecall\_one\text{-}month} + B_{3ELL} + B_{4Question\_Type} + \varepsilon_i$

*Note: Pretest is grand-mean centered.

Results 1.3

Table 1 provides demographic information including students' average scores on reading comprehension and fluency. As shown in Table 1, there were no statistically significant differences on the GMRT or FAST between ELs and monolingual English speakers. This is important because the purpose of this study was to examine knowledge-based inferencing differences among ELs and monolingual English-speaking students who are typically developing fluent readers, where typical is defined as adequate reading comprehension and fluent is defined as reading fluency performance above state benchmarks. Additionally, as shown in Table 1, results of a chi-square goodness of fit test showed no association between gender and grade. ELs were predominantly from Asian or Hispanic/Latino ethnicities while monolingual English speakers were mostly Caucasians.

(1) Are there differences in learning and retention of information from a novel knowledge base between ELs and monolingual English speakers?

Table 2 reports students' performance on the ease of learning the newly taught knowledge base as measured by the forced choice picture recognition task and verbal recall task. Results show that although ELs and monolingual English speakers performed comparably on the forced choice picture recognition task (g = -0.15, p > .05), ELs required significantly more trials to achieve perfect recall of the 14 facts about the planet Gan (g = -0.42, p < .05) as measured by the verbal recall task. Scores on both measures suggest that, as intended, both ELs and

monolingual English-speaking students acquired the knowledge base with relative ease. However, compared to monolingual English speakers, ELs acquired the knowledge base through more repetitions of the content.

When measuring students' stability and retention of the taught knowledge base over time, ELs and monolingual English speakers did not significantly differ immediately or at the one-week and one-month follow-up time points. Across the three time points, effect sizes ranged from $g = 0.17$ to $g = 0.33$ in favor of monolingual English speakers. Effect size differences on the verbal recall task indicate that monolingual English speakers retained and recalled the knowledge base at higher mean levels of accuracy which may be practically significant albeit not statistically significant.

(2) How do coherence and elaborative inferences compare in English Learners and monolingual English students after learning the novel knowledge base? (3) Does inference making in the context of text comprehension differ from inference making elicited after the story concludes?

To examine whether differences existed between ELs' and monolingual English speakers' ability to accurately form coherence and elaborative inferences, scores were analyzed using regression models that controlled for student group (EL, monolingual English speaker), inference type (coherence, elaborative), and access to knowledge base (i.e., verbal recall) at each of the three time points (immediate, one-week follow-up, and one-month follow-up). Student group represented a between subject factor and inference type and access to knowledge base within subject factors. Table 3 indicates that the mean inference-making scores of ELs were consistently lower relative to monolingual English-speakers across each time point. This pattern of findings was consistent for both coherence and elaborative inference question types.

Specifically, the magnitude of difference in coherence inference-making between ELs and monolingual English speakers ranged from $g = 0.31$ to $g = 0.55$ in favor of the monolingual English speakers across phases two and three testing time points. The magnitude of the difference in elaborative inference making ranged from $g = 0.35$ to $g = 0.41$, also in favor of monolingual English speakers.

Table 4 reports whether performance on inference questions varied by students' performance on the verbal recall task and the type of inference question after controlling for EL status and their pretest reading comprehension scores. Results demonstrate that students' performance on the verbal recall task at each time point was a significant predictor of their inference score. In Table 4, models 1a, 1b, and 1c demonstrate that when controlling for students' EL status and pretest reading comprehension scores, higher scores on verbal response task at each time point was associated with greater accuracy on the knowledge based inference measure immediately (B =1.50, SE =.33, $p < .01$), at the one-week follow-up (B =0.90, SE =.09, $p < .01$), and at the one-month follow-up (B =0.83, SE =.06, $p < .01$) time points.

Additionally, in models 2a, 2b, and 2c, as shown in Table 4, we controlled for the type of inference question (i.e., elaborative or coherence). When controlling for EL status, pretest reading comprehension, and verbal recall accuracy scores, students performed significantly better on elaborative inference questions compared to coherence inference questions immediately (B =0.45, SE =.17, $p < .05$) and at one-week follow-up (B =0.46, SE =.16, $p < .05$). There was no significant difference in accuracy on the two inference question types at the one-month follow-up (B =0.20, SE =.12, $p >.05$).

Furthermore, Models 1a, 1b, and 1c in Table 4 demonstrate that after controlling for students' verbal recall accuracy and pretest reading comprehension scores, ELs performed

significantly below monolingual English-speakers immediately (*B* = -0.88, *SE* =.41, *p* >.05), one-week (*B* = -0.92, *SE* =.38, *p* <.05), and one-month (*B* = -0.66, *SE* =.28, *p* <.05) follow-up time points on the knowledge base inference measure. As shown in Models 2a, 2b, and 2c, results were similar when controlling for the type of inference question immediately (*B* = -0.46, *SE* =.17, *p* <.05), at the one-week (*B* = -0.51, *SE* =.16, *p* <.05), and at the one-month (*B* = -0.35, *SE* =.12, *p* <.05) follow-up time points; that is, ELs performed significantly worse compared to monolingual English speakers. Finally, we tested if there were significant interactions between EL status and inference question type at each of the three time points. There were no significant interactions between EL status and questions type at the immediate (*B* = 0.17, *SE* =.35, *p>.05),* one-week (*B* = 0.08, *SE* = 0.33, *p>.05),* and one-month (*B* = 0.21, *SE* =.25, *p>.05)* follow-up time point; to present a parsimonious model, we dropped the interaction term from the final models in Table 4.

Discussion 1.4

Findings from this study describe the knowledge-based inference making skills of English learners and monolingual English speakers in the middle grades where a new knowledge base was taught to all students and where the only inferences required were those that drew on this newly learned knowledge base. The use of a newly learned knowledge base provided the opportunity to examine how knowledge availability is related to coherence and elaborative inference making among ELs, the fastest growing subgroup of students attending United States schools. To accurately make either type of inference, evidence from text and information from the taught knowledge base had to be activated, retrieved, and integrated. We summarize and discuss the results below.

Learning and retention of information from a novel knowledge-base 1.4.1

First, the experimental paradigm was designed to ensure that both ELs and monolingual English speakers learned a set of knowledge base elements from which inferences could be made with relative ease. To accomplish this, the 14 facts about the planet "Gan" were taught to perfect recall, resulting in ceiling effects for both groups. Results show that ELs and monolingual English-speakers learned the novel knowledge base quickly. However, compared to monolingual English speakers, ELs required slightly more learning trials to achieve perfect recall. More specifically, the effect size of .42 indicates that ELs required approximately one additional learning trial to perfectly recall all 14 knowledge base elements about the planet "Gan." This difference in learning was statistically significant, suggesting that the learning rate differences between monolingual English speakers and ELs are practically meaningful. These findings also provide practical suggestions for educators in that equivalent outcomes for ELs and monolingual English students were possible when ELs were provided immediate feedback for all wrong answers and incorrect items were practiced until perfectly recalled.

Stability and retention of the newly taught knowledge base was assessed immediately after listening to the seven-episode story, one week and one month later. Stability of the knowledge base was re-assessed to determine whether inference skill was associated with retention of the knowledge base over time and whether group differences in the stability of the knowledge base existed. Across testing time points, ELs and monolingual English speakers recalled the knowledge base at similar levels of accuracy, suggesting that the knowledge base was equally available over the time students were required to draw on this knowledge to make inferences. Consequently, any differences in inferencing between ELs and monolingual English-speaking students cannot be attributed to variations in the availability of knowledge about "Gan."

Coherence and elaborative inferencing among ELs and monolingual English speakers 1.4.2

Second, the experimental design enabled us to describe elaborative and coherence inferencing among ELs and monolingual English-speaking students when knowledge was equally available to both groups of students. In addition, students' inference-making accuracy was assessed in the context of story comprehension as well as immediately afterwards, one week, and one month later. These three delayed assessments of inference-making allowed us to compare inference-making among ELs and monolingual English-speaking students when inference-making was cued to evaluate if inferencing difficulties signify more fundamental reasoning and integration issues involved in generating coherence and elaborative inferences (Barnes et al., 1996).

Results show that even though ELs learned the knowledge base and retained the same amount of information over time, they generated significantly fewer elaborative and coherence inferences than monolingual English-speaking students. The effect size differences indicate that differences in coherence and elaborative inference-making between ELs and monolinguals were substantial. So, even though ELs had acquired the knowledge necessary to make these two types of inferences, they were not drawing on the taught knowledge base as consistently as monolingual English-speaking students.

Next, comparing the accuracy by which elaborative and coherence inferences were made immediately after the story, we see that for every additional knowledgebase item that was accurately recalled, inferencing increased by 1.5 points among both groups. Similarly, at the one week and one month follow up time points, we see that for every additional knowledge base element recalled accurately, inferencing improved by 0.83-0.90 points. Results suggest that knowledge-based inferences are constructed to the extent that the information on which the

inference depends is readily available. If required information is readily available, then the inference is more likely to be accurately made.

In addition, the study compared knowledge-based inferencing for two reading contexts. In the first context, students formed inferences during text comprehension. This context can be considered a relatively complex processing condition due to the amount of information that must be processed to understand the story. In the second condition, students were asked to form knowledge-based inferences immediately after the story, one week, and one month later. This condition is less complex because the questions cued textual information that should be integrated with the knowledge base to form the inference. If ELs are poorer at integrating information to form knowledge-based inferences even in simple processing contexts, then their inferencing difficulties during text comprehension may reflect basic limitations in the operation of integrating information or reasoning and not the operation of inferencing per se (Barnes, 1996). Results revealed that although monolingual English-speaking students outperformed ELs in the formation of knowledge-based inferences, both groups of students generated a greater number of coherence inferences compared to elaborative inferences while listening to the story episode. At immediate post-test and one week follow-up, students generate significantly more elaborative inferences. Finally, at the one month follow-up, students generated more coherence than elaborative inferences even though the means were practically similar.

Our findings are similar to that of past research with typically developing children which shows that coherence inferences were made during text comprehension but more elaborative inferences were made in the less complex processing condition (immediate and one week follow-up) where the question cued the inference. Results align with past research showing that coherence inferences were made more often than elaborative inferences during ongoing

comprehension of text (Barnes et al., 1996). The idea that the integration or reasoning skills required for elaborative inferencing are simpler than those skills required for coherence inferencing is consistent with the notion that while elaborative inferences are made less often than coherence inferences during text comprehension, they can be easily made from a situation model of the text (Garham, 1982). Further, this study also suggests that coherence inferences are often made during text comprehension to support understanding at the local level since texts rarely provide all the details of an event required to fully describe the situation in a lifelike manner (Oakhill & Garnham, 1981; Barnes et al., 1996). This implies that coherence inferences may be obligatory such that a text cannot be fully understood without them (Gerrig, 1986). Of note, this study is the first to show that coherence inferences are regularly encoded by ELs during reading when information on which the inference depends is readily available, thus expanding our understanding of constructionist models of text comprehension.

Finally, because the experimental design first taught the knowledge base to mastery then prompted students to form inferences using this knowledge, we can argue that knowledge availability is essential for the accurate generation of knowledge-based inferences. Of note, by the one month follow-up, after controlling for student EL status and their pretest comprehension score, knowledge availability accounted for 60% of the variance in knowledge-based inferencing. However, the availability of relevant knowledge did not facilitate knowledge-based inferencing among ELs at the same level of accuracy as monolingual English-speaking students. This suggests that for students learning English as a second language, immediately available information from semantic memory may be more varied than we currently believe and that knowledge available in short-term memory may be more complexly structured than originally thought (McKoon & Ratcliff, 1992). Perhaps a contributing factor could be that the richness of

the knowledge base or depth of concepts used in the "Gan" task is not completely equivalent between the two groups even though ELs were trained to accurate recall. Alternately, the organization of the "Gan" knowledgebase may not be equivalent thus impacting lexical retrieval accuracy and lexical retrieval rates among ELs. That is, weaker links between existing general knowledge and associated concepts in long-term memory, potentially due to less accumulated practice with the English language, may lead to less accurate and slower retrieval of newly acquired knowledge and greater inferencing errors among ELs.

To summarize, although the importance of general knowledge in the construction of inferences is well established among monolingual English-speaking students, the present study substantiates its role among ELs. Specifically, this study shows that knowledge is a necessary but not sufficient condition for inferencing among monolingual English-speaking students and ELs in the rural, middle grades. Further, this study indicates that knowledge-based inferencing is an area for providing additional instructional support for ELs and that their inference difficulties are not wholly accounted for by the availability of general knowledge.

Study Limitations 1.4.3

We recognize several limitations of the present study. First, the sample was relatively small which may partly explain why interaction effects were null. Future research should be conducted on a larger scale so there is greater power to detect effects with randomized control designs and to explore how effects might differ as a function of different learning conditions and types of inferences. Second, results of this study are based on a single sample of students from one rural school district who were identified using only one reading comprehension and one fluency measure. More studies of this nature are needed to generalize the current study's findings. Third, the study was conducted among ELs who performed above school district

reading benchmarks. English learners with reading difficulties were not included. This population of ELs were the focus of this study given their growing proficiency in English but access to limited supplemental support within general education middle grade classrooms (Proctor et al., 2019), particularly classrooms in rural districts. Further, general education teachers generally agree that strategies to improve comprehension, such as inferencing, would be useful to ELs who are still learning to use their linguistic resources to independently learn from content area texts (Proctor et al., 2019). However, because the majority of ELs read below proficient levels, an important next step is to examine inference making skills among ELs who read below grade level benchmarks and ELs identified with a reading disability as well as the effect of knowledge retrieval on inference-making (Hall et al., 2017).

Conclusions 1.5

We believe that our findings have implications for future research and practice. First, when typically developing ELs were provided sufficient learning trials, they remembered new knowledge at the same level of accuracy as monolingual English-speaking students. Yet, mere availability of knowledge in semantic memory did not ensure that such knowledge was accessed and integrated to form coherence and elaborative inferences. These results remind us that learning facts are important for thinking, reasoning, and problem solving, but expertise depends significantly on a rich body of knowledge about content matter (e.g., Chase & Simon, 1973; Chi et al., 1981; DeGroot, 1965). In other words, "usable knowledge" for the purpose of thinking, reasoning, and problem solving is not the same as learning and recalling a list of disconnected facts about specific content. Experts' knowledge is connected and organized around important concepts that is both readily available and accessible for the specific contexts for which it is appropriate to apply; it supports understanding and learning as well as transfer to other contexts

(National Research Council, 2000). For this reason, it is important to consider how to help ELs build knowledge structures that are smartly organized and well-connected to and overlapping with related concepts in long term memory.

Second, this study extends previous research (e.g., Barnes & Dennis, 1998; Barnes et al., 1996; Author, 2001) by showing that typically developing ELs do not make knowledge-based inferences at the same level of accuracy as monolingual English-speaking students. These findings are important and suggest that inference-making, a higher-level language skill, may represent a key instructional target that will enable higher levels of comprehension for ELs. Results also provide support for Proctor and colleagues' claim that to improve typically developing ELs access to and understanding of the curriculum, instruction in the general education classroom may require broader emphasis on dimensions of language beyond academic vocabulary and reading comprehension generally. Yet, for typically developing ELs, inference-making has largely been overlooked as an instructional target (Hall et al., 2017; Proctor et al., 2019).

Acknowledgements 1.6

The authors wish to acknowledge the invaluable contributions of the middle school staff and students in Storm Lake Middle School, Storm Lake Community School District, and the assistance of Principal Jay Slight, Assistant Principal Jamie Braal, and Instructional Coach Megan Richardson.

Figure 1

Study design.

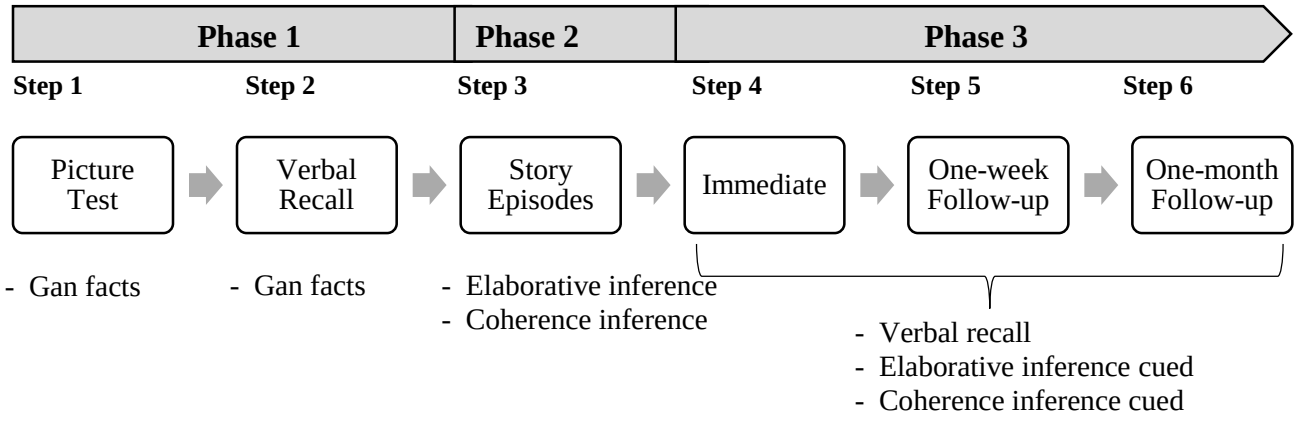


Table 1

Demographic Information

Measure	ELs (n = 87)	Non-ELs (n = 58)	χ^2	*p*-value
Gender				
- Male	38	23	0.095	.757
- Female	49	35		
Grade				
- Fifth	54	33	0.202	.652
- Sixth	33	25		
Ethnicity / Race				
- Asian	22	1	82.71	<.01
- African American	3	2		
- Caucasian	0	37		
- Hispanic / Latino	59	16		
- Other	3	2		
GMRT – Reading Std Score				
- Fifth Grade	102.29 (7.96)	103.38 (8.35)	-	.544
- Sixth Grade	103.66 (11.44)	108.74 (10.02)	-	.083
- Full Sample	102.81 (9.39)	105.69 (9.41)	-	.072
CBM WCPM				
- Fifth Grade	163.92 (25.30)	162.93 (24.73)	-	.858
- Sixth Grade	171.53 (16.46)	174.12 (20.04)	-	.591
- Full Sample	166.78 (22.58)	168.01 (23.20)	-	.751

Note: EL – English language learner. Non-Els – Monolingual English speakers. GMRT Reading = Gates MacGinitie Reading Comprehension Test; CBM WCPM = Fast Bridge CBM Reading Words Correct Per Minute

Table 2

ELs and Non-ELs Ease of Learning and Retention of the Knowledge Base during Phase 1 and Phase 3

	Ease of Learning (Phase 1)				Retention (Phase 3)					
	Picture Test		Verbal Recall		Immediate (Verbal Recall)		1-week delay (Verbal Recall)		1-month delay (Verbal Recall)	
EL Status	M	SD	M	SD	M	SD	M	SD	M	SD
Non-EL	15.57	1.84	15.69	2.14	13.82	1.87	13.72	4.00	12.47	2.73
EL	15.90	2.40	16.66	2.44	13.49	0.59	12.74	1.94	12.04	2.29
Hedge's *g*	-0.15		-0.42*		0.26		0.33		0.17	

EL – English language learner; Non-Els – Monolingual English speakers

Note—One point was given for each item recalled accurately for the first time on the picture test and verbal recall test. For inaccurate responses, errors were corrected and retested, and 1 point was added for additional trials needed. Perfect recall score was 14. Higher scores on ease of learning tasks indicate greater number of errors. For the immediate and delayed tests of recall, the maximum score was 14 too. Higher scores on retention tests indicate greater retention.

*p<.05

Table 3

Average Coherence and Elaborative Inferencing Scores for ELs and Non-ELs

Phase	Inferencing Type	Max Score Possible	Non-ELs	ELs	Hedge's *g*
	Story Episodes				
2	Coherence	14	7.75 (2.90)	6.27 (2.47)	0.55*
	Elaborative	14	6.15 (2.43)	5.31 (2.20)	0.36 †
	Immediate				
3	Coherence	7	5.52 (1.43)	4.75 (1.75)	0.47*
	Elaborative	7	5.78 (1.36)	5.28 (1.45)	0.35†
	One-week Follow-up				
3	Coherence	7	5.63 (1.39)	4.97 (1.80)	0.37†
	Elaborative	7	5.98 (1.23)	5.38 (1.57)	0.41†
	One-month Follow-up				
3	Coherence	7	5.95 (1.22)	5.51 (1.55)	0.31
	Elaborative	7	5.86 (1.38)	5.37 (1.42)	0.35†

*Note: In Phase 2, two points were recorded for unprompted correct response, one point for prompted correct response, and no points for incorrect response. In Phase 3, one point was recorded for unprompted correct response and no points for incorrect response; students were not prompted during Phase 3.

EL – English language learner; Non-Els – Monolingual English speakers

**p*<.05 (after controlling for the false discovery rate using the Benjamini-Hochberg procedure)

† p-value <.05 (prior to Benjamini-Hochberg correction)

Table 4

Regression Models Predicting Student Response to Know-it Inference Task during Phase 3

	Model 1a	Model 2a	Model 1b	Model 2b	Model 1c	Model 2c
Predictors	*Estimates* (*SE*)	*Estimates* (*SE*)	*Estimates* (*SE*)	*Estimates* (*SE*)	*Estimates* (*SE*)	*Estimates* (*SE*)
Intercept	-9.26† (4.56)	1.18 (1.33)	-0.12 (1.30)	0.27 (0.52)	1.63† (0.80)	1.14** (0.33)
Verbal Recall Immediate	1.50** (0.33)	0.60** (0.19)				
Verbal Recall One-Week			0.90** (0.09)	0.81** (0.07)		
Verbal Recall One-Month					0.83** (0.06)	0.76** (0.05)
EL–Yes	-0.88† (0.41)	-0.46* (0.17)	-0.92* (0.38)	-0.51** (0.16)	-0.66* (0.28)	-0.35* (0.12)
GMRT	0.07** (0.02)	0.04** (0.00)	0.03 (0.02)	0.02* (0.00)	0.03 (0.01)	0.02* (0.00)
Question Type - Elaborative		0.45* (0.17)		0.46* (0.16)		0.20 (0.12)
R^2	.265	.170	.477	.398	.598	.498

Note: Note: Models a = Immediate posttest; Models b = One-week follow-up; Models c = One-month follow-up. EL = English Learner. GMRT = Gates MacGinitie Reading Comprehension Pretest (Grand-mean centered).

***p*<.01, **p*<.05 (after controlling for the false discovery rate using the Benjamini-Hochberg procedure)

† p-value <.05 (prior to Benjamini-Hochberg correction)

References

Abu-Rabia, S. (1996). Factors affecting the learning of English as a second language in Israel. *The Journal of Social Psychology, 136*(5), 589–595.

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M., & Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7 through 12. *Contemporary Educational Psychology, 44*, 68–82. https://doi.org/10.1016/j.cedpsych.2016.02.002

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(5), 1061.

August, D., Francis, D. J., Hsu, H. Y. A., & Snow, C. E. (2006). Assessing reading comprehension in bilinguals. *The Elementary School Journal, 107*(2), 221–238.

August, D., & Shanahan, T. (2006). Synthesis: Instruction and professional development. *Developing Literacy in a Second Language: Report of the National Literacy Panel*, 321–335.

Barnes, M. A., Ahmed, Y., Barth, A., & Francis, D. J. (2015). The relation of knowledge-text integration processes and reading comprehension in 7th- to 12th-grade students. *Scientific Studies of Reading, 19*(4), 253–272. https://doi.org/10.1080/10888438.2015.1022650

Barnes, M. A., & Dennis, M. (1998). Discourse after early-onset hydrocephalus: Core deficits in children of average intelligence. *Brain and Language, 61*(3), 309–334. https://doi.org/10.1006/brln.1998.1843

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from

six to fifteen years of age. *Journal of Experimental Child Psychology*, *61*(3), 216–241. https://doi.org/10.1006/jecp.1996.0015

Barth, A. E., Barnes, M., Francis, D., Vaughn, S., & York, M. (2015). Inferential processing among adequate and struggling adolescent comprehenders and relations to reading comprehension. *Reading and Writing*, *28*(5), 587–609. https://doi.org/10.1007/s11145-014-9540-1

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & cognition*, *29*(6), 850–859. https://doi.org/10.3758/bf03196414

Hall, C., Roberts, G. J., Cho, E., McCulley, L. V., Carroll, M., & Vaughn, S. (2017). Reading instruction for English learners in the middle grades: A meta-analysis. *Educational Psychology Review*, *29*(4), 763-794. https://doi.org/10.1007/s10648-016-9372-4

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. doi:10.1111/j.2517-6161.1995.tb02031.x

Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics*, *25*(1), 60–83.

Bloom, C. P., Fletcher, C. R., Van Den Broek, P., Reitz, L., & Shapiro, B. P. (1990). An on-line assessment of causal reasoning during comprehension. *Memory & Cognition*, *18*(1), 65-71.

Bonifacci, P., Canducci, E., Gravagna, G., & Palladino, P. (2017). English as a foreign language in bilingual language-minority children, children with Dyslexia and monolingual typical readers. *Dyslexia*, *23*(2), 181–206.

Bonifacci, P., & Tobia, V. (2016). Crossing barriers: Profiles of reading and comprehension skills in early and late bilinguals, poor comprehenders, reading impaired, and typically developing children. *Learning and Individual Differences*, *47*, 17–26.

Bransford, J. D., Stein, B. S., Shelton, T. S., & Owings, R. A. (1981). Cognition and adaption: The importance of learning to learn. In J. Harvey (Ed.), *Cognition, social behavior and the environment*. Hillsdale, N.J.:Earlbaum

Bråten, I., & Samuelstuen, M. S. (2004). Does the Influence of Reading Purpose on Reports of Strategic Text Processing Depend on Students' Topic Knowledge? *Journal of Educational Psychology*, *96*(2), 324–336. https://doi.org/10.1037/0022-0663.96.2.324

Burgoyne, K., Whiteley, H. E., & Hutchinson, J. M. (2013). The role of background knowledge in text comprehension for children learning English as an additional language. *Journal of Research in Reading*, *36*(2), 132–148.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., ... & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, *39*(2), 188–215.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, *18*(3), 257–273.

Christ, T.J. (2015). Formative Assessment System for Teachers: Technical Manual Version 2.0, Minneapolis, MN: Author and FastBridge Learning (www.fastbridge.org).

Compton, D. L., Miller, A. C., Elleman, A. M., & Steacy, L. M. (2014). Have we forsaken reading theory in the name of "quick fix" interventions for children with reading disability? *Scientific Studies of Reading*, *18*(1), 55–73.

Currie, N. K., & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology*, *137*, 57-75.

D'Angelo, N., & Chen, X. (2017). Language profiles of poor comprehenders in English and French. *Journal of Research in Reading*, *40*(2), 153–168.

Daniel, J., Capin, P., & Steinle, P. (2020). A synthesis of the sustainability of remedial reading intervention effects for struggling adolescent readers. *Journal of Learning Disabilities*. Advance online publication. https://doi.org/10.1177/0022219420972184

DeLeon, P. H., Wakefield, M., & Hagglund, K. J. (2003). The behavioral health care needs of rural communities. In B. H. Stamm (Ed.), *Rural behavioral health care: An interdisciplinary guide* (p. 23–31). American Psychological Association. https://doi.org/10.1037/10489-001

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197. http://dx.doi.org/10.1017/S1366728902003012

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading*, *17*(6), 435–452.

Francis, W. S., Strobach, E. N., Penalver, R. M., Martínez, M., Gurrola, B. V., & Soltero, A. (2019). Word–context associations in episodic memory are learned at the conceptual

level: Word frequency, bilingual proficiency, and bilingual status effects on source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(10), 1852–1871. https://doi.org/10.1037/xlm0000678

Galloway, E. P., & Lesaux, N. (2015). Reading comprehension skill development and instruction for adolescent English language learners: A focus on academic vocabulary instruction. In *Improving reading comprehension of middle and high school students* (pp. 153-178). Springer, Cham.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 430.

Geva, E., & Massey-Garrison, A. (2013). A comparison of the language skills of ELLs and monolinguals who are poor decoders, poor comprehenders, or normal readers. *Journal of Learning Disabilities*, *46*(5), 387–401.

Goldenberg, C. (2011). Reading instruction for English language learners. In M. Kamil, P. D. Pearson, E. Moje, & P. Afflerbach (Eds.), *Handbook of Reading Research, Volume IV*, (pp. 684–710). Newark, DE: International Reading Association.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*,787–814. https://dx.doi.org/10.1016/j.jml.2007.07.001

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67–81. http://dx.doi.org/10.1017/S1366728998000133

Haenggi, D., & Perfetti, C. A. (1992). Individual differences in reprocessing of text. *Journal of Educational Psychology*, *84*(2), 182–192. https://doi.org/10.1037/0022-0663.84.2.182

Hoff E, Core C, Place S, Rumiche R, Señor M, Parra M. Dual language exposure and early bilingual development. Journal of Child Language. 2012; 39:1–27. DOI: 10.1017/S0305000910000759 [PubMed: 21418730]

Kembo, J. A. (2001). Testing of inferencing behaviour in a second language. *International Journal of Bilingual Education and Bilingualism*, *4*(2), 77–96.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., ... & Velez, E. D. V. (2015). The condition of education 2015. NCES 2015-144. *National Center for Education Statistics*.

Kendeou, P. (2015). A general inference skill. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp.160–181). Cambridge, UK: Cambridge University Press.

Kendeou, P., & Van Den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, *97*(2), 235–245.

Kendeou, P., & Van Den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, *35*(7), 1567–1577.

Kieffer, M. J. (2010). Socioeconomic status, English proficiency, and late-emerging reading difficulties. *Educational Researcher*, *39*(6), 484–486.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182.

Kintsch, W. (1993). Information accretion and reduction in text processing: Inferences. *Discourse Processes, 16*(1-2), 193-202.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*(4), 294.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174. http://dx.doi.org/10.1006/jmla.1994.1008

Lesaux, N. K., & Geva, E. (2006). Synthesis: Development of Literacy in Language-Minority Students. In D. L. August & T. Shanahan (Eds.), *Developing Literacy in a Second Language: Report of the National Literacy Panel* (pp. 5374).

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*, *45*(2), 196–228.

Li, M. (2012). *Understanding reading comprehension in English immersion students in China* (Unpublished doctoral dissertation). Queen's University.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2000). *Gates-MacGinitie Reading Tests* (fourth ed.). Iasca: Riverside.

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology*, *102*(3), 701–711.

Mancilla-Martinez J, Vagh, SB. Growth in toddlers' Spanish, English, and conceptual vocabulary knowledge. Early Childhood Research Quarterly. 2013; 28:555–567. https://doi.org/0.1016/j.ecresq.2013.03.004

Martinez-Lincoln, A., Barnes, M.A., Clemens, N., & Tran, L. (2021, June). *Inference-making in students with diverse linguistic backgrounds: A systematic review.* Oral presentation and poster presented at the Extraordinary Brain Symposium XVIII Meeting: Dyslexia in Vulnerable Student Populations Student Populations-Leading for Change, Guayaquil, Ecuador. Canceled due to COVID-19.

McKeown, M., Beck, I., Sinatra, G., & Loxterman, J. (1992). The Contribution of Prior Knowledge and Coherent Text to Comprehension. *Reading Research Quarterly, 27*(1), 79-93. https://doi.org/10.2307/747834

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*(3), 440.

McNamara, D. S., de Vega, M., & O'Reilly, T. (2007). Comprehension skill, inference making, and the role of knowledge. In F. Schmalhofer & C.A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 233–251). Mahwah, NJ: Erlbaum.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*(3), 247-288.

Miller, A. C., & Keenan, J. M. (2009). How word decoding skill impacts text memory: The centrality deficit and how domain knowledge can compensate. *Annals of Dyslexia, 59*(2), 99.

Miller, A. C., & Keenan, J. M. (2011). Understanding the centrality deficit: Insight from foreign language learners. *Memory and Cognition, 39,* 873–883. https://doi.org/10.3758/s13421-010-0062-z

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2008). A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading*, *12*(4), 351–371.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology*, *54*(1), 31–39.

Oh, S. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly*, 35, 69-96. https://doi.org/10.2307/3587860

Prior, A., Goldina, A., Shany, M., Geva, E., & Katzir, T. (2014). Lexical inference in L2: Predictive roles of vocabulary knowledge and reading skill beyond reading comprehension. *Reading and Writing*, *27*(8), 1467–1484.

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, *97*(2), 246–256.

Proctor, C. P., Silverman, R. D., Harring, J. R., Jones, R. L., & Hartranft, A. M. (2020). Teaching bilingual learners: Effects of a language- based reading intervention on academic language and reading comprehension in grades 4 and 5. *Reading Research Quarterly*, *55*(1), 95-122.

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from https://www.pstnet.com.

R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rawson, K. A., & Kintsch, W. (2004). Exploring encoding and retrieval effects of background information on text memory. *Discourse processes, 38*(3), 323-344.

Rawson, K. A., & Van Overschelde, J. P. (2008). How does knowledge promote memory? The distinctiveness theory of skilled memory. *Journal of Memory and Language, 58*(3), 646–668.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, *80*(1), 16–20.

Reed, D. K., Petscher, Y., & Foorman, B. R. (2016). The contribution of vocabulary knowledge and spelling to the reading comprehension of adolescents who are and are not English language learners. *Reading and Writing*, *29*(4), 633–657.

Ryan, C. L. (2013). Language use in the United States: 2011. Washington, DC: U.S. Census Bureau.

Schneider, W., Eschman, A., and Zuccolotto, A. (2012). E-Prime Reference Guide. Pittsburgh: Psychology Software Tools, Inc.

Snijders, T. A., & Bosker, R. J. (1999). *An introduction to basic and advanced multilevel modeling*. Sage, London.

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.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2017 Mathematics and Reading Assessments. Retrieved from the Nation's Report Card, March, 2020.

Van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse processes*, *39*(2-3), 299–316.

Van Hell, J. G., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning, 47,* 507–546. http://dx.doi.org/10.1111/0023-8333.00018

Verhoeven, L. (2011). 27 Second Language Reading Acquisition. In M. Kamil, P. D. Pearson, E. Moje, & P. Afflerbach (Eds.), *Handbook of Reading Research, Volume IV*, (661–683). Newark, DE: International Reading Association.

Appendix A: Knowledge Base

The rivers and ponds on Gan are filled with orange juice!

The trees on Gan have pink leaves!

The shoes from Gan have wings!

The mushrooms on Gan are as tall as the children!

The bears on Gan have blue fur!

The raindrops on Gan are shaped like triangles!

The cats on Gan have to be dressed in a coat and a hat and shoes before they can go outside!

The mountains on Gan are made of bubbles!

The walls and roofs of the houses on Gan are built of glass!

The frogs on Gan glow in the dark!

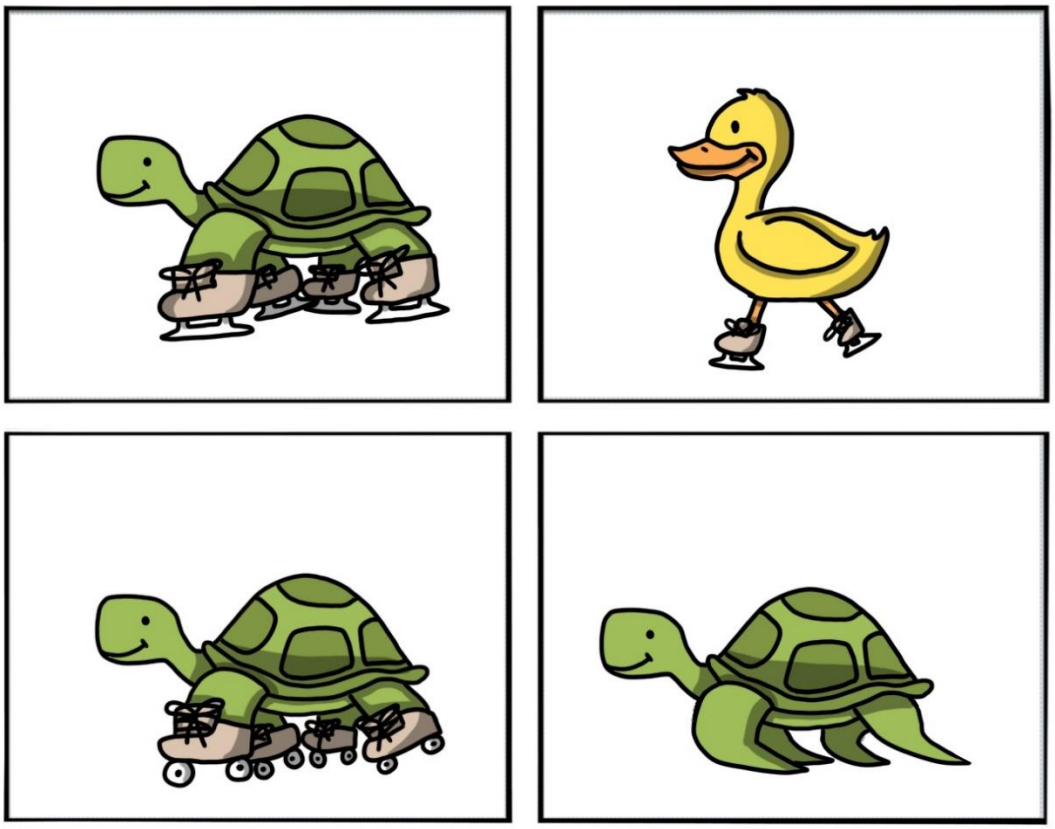
The book covers on Gan are made of popcorn!

On Gan, all the people have bright green hair!

The turtles on Gan wear ice skates!

The moon on Gan is shaped like a diamond!

Appendix B: Force-choice picture-selection task



Appendix C: Example of Story Episode and Questions

Episode 1

Dack was lying on his bed under the attic roof. He looked up and saw a squirrel enjoying the sun and eating an acorn on the roof. “What should we do today?” he yawned. “Dack, did you forget? Our cousin is coming to visit,” said Tane. Dack and Tane decided to run away from home because their cousin always got them in trouble. Dack and Tane packed their backpacks and were just about to leave when their cat, Zoe, spied them and began to meow loudly. “Sshhh! Dack whispered. “You’ll wake mom and dad.” Tane got the cat ready to go. Then the children took their shoes off and tiptoed outside.

Questions

1. What did Dack see when he woke up? (coherence inference)
2. What did Tane do to Zoe before leaving? (elaborative inference)

Appendix D: Cued Coherence and Elaborative Inference Task

Inference Question	Inference Type
1. How could Dack see the squirrel eating on the roof?	Coherence
2. How did Zoe look when Tane got her ready to go?	Elaborative
3. How did Dack and Tane manage to fly up and over the fence?	Coherence
4. What did Dack and Tane have to drink when they drank from the pond?	Elaborative
5. How could Dack and Tane blow pieces of the mountain at each other?	Coherence
6. What colour was the ribbon that Tane used to match her hair?	Elaborative
7. Why did Dack and Tane run to the mushroom when it started to rain?	Coherence
8. When Dack watched the rain fall, what did it look like?	Elaborative
9. Why did Dack wish he was a turtle?	Coherence
10. What colour were the bear fur coats that Dack and Tane took out of their backpacks?	Elaborative
11. Why did Dack and Tane use frogs to help them find a cave in the dark?	Coherence
12. When Dack looked up at the moon, what did it look like?	Elaborative
13. How could Dack and Tane eat the covers of their storybooks for a snack?	Coherence
14. What colour were the beds of leaves that Dack and Tane made?	Elaborative

Appendix E: Information on Excluded Sample

Reason for exclusion	Grade 5		Grade 6	
	EL (n = 57)	Non-EL (n = 17)	EL (n = 38)	Non-El (n = 13)
GMRT < 85	19	7	13	8
Below WCPM Benchmark*	21	9	16	4
GMRT < 85 and Below WCPM Benchmark*	17	1	9	1

* GMRT = Gates MacGinitie Reading Comprehension Pretest; WCPM = FastBridge Oral Reading Fluency Words Correct Per Minute.

Note: Grade 5 CBM Benchmark <133 words correct per minute; Grade 6 CBM Benchmark < 142 words correct per minute.