**The influence of reader and text characteristics on sixth graders' inference making**

Prof Kate Cain*, Lancaster University, UK. k.cain@lancaster.ac.uk

Dr Nicola K. Currie, Edge Hill University, UK. Nicola.Currie@edgehill.ac.uk

Dr Gillian Francey, Lancaster University, UK. g.francey@lancaster.ac.uk

Dr Robert Davies, Lancaster University, UK. r.davies1@lancaster.ac.uk

Dr. Shelley Gray, Arizona State University, U.S. Shelley.Gray@asu.edu

Dr. Mindy S. Bridges, University of Kansas Medical Center, U.S. mbridges2@kumc.edu

Dr M. Adelaida Restrepo, University of South Florida, U.S. marestrepo@usf.edu

Dr. Marilyn S. Thompson, Oregon State University, U.S.

Marilyn.Thompson@oregonstate.edu

Dr Margeaux F. Ciraolo, Arizona State University, U.S. margeaux.ciraolo@asu.edu

**Author Note**

*corresponding author

Our pre-registered data preparation and analysis plan can be found at (https://osf.io/j7zps).

The inference texts, dataset and analysis code are deposited at (https://osf.io/xs2um/).

The study conforms to the US Federal Policy for the Protection of Human Subjects and was approved by the Institutional Review Board or Research Ethics Committee at each university.

Accepted for publication in Journal of Research in Reading, 08/10/2024

## Abstract

**Background:** We examined the influence of text and reader characteristics on sixth graders' inference generation.

**Methods:** Eleven to 12-year-old U.S. monolingual English speakers (N=71) and Spanish-English bilinguals (N=81) read narrative and informational expository texts requiring an inference and answered an inference-tapping question after each text. We examined the influence of language status, word reading ability, knowledge (background knowledge and vocabulary), and reading strategy awareness and use on question accuracy, question-answering times, and sentence reading times.

**Results:** Linear mixed effects models predicting response accuracy indicated an advantage for narrative texts, in general, and for participants with higher knowledge. When examining variation across the whole sample, rather than contrasting language groups, faster question-answering and sentence processing times were associated with higher knowledge.

**Conclusions:** Adolescent readers are better able to generate inferences from narrative than informational expository texts, and knowledge has a critical influence on both the process and product of inference generation and may explain reading comprehension performance differences between monolingual and bilingual students.


**Keywords:** Inference; reading comprehension; narrative text; expository text; knowledge; adolescent students.

**Implications for Practice**

**What is already known about this topic**

- Inference making is critical for adequate comprehension of narrative and informational expository texts.
- Background knowledge and reading strategies are both associated with comprehension of inferential texts.
- Monolingual and bilingual readers differ in knowledge and reading comprehension performance.

**What this paper adds**

- Adolescent readers are better able to generate inferences from narrative than expository texts.
- Individual differences in knowledge explains performance in both accuracy of inference making and inferential processing.
- Knowledge differences between monolingual and bilingual readers may explain the poorer inference performance by bilingual students.

**Implications for theory, policy or practice**

- Instruction in the acquisition of knowledge across topics would support inference generation from both narrative and expository texts.
- Knowledge should be considered as a potential source of weak inference making and reading comprehension.

**The influence of reader and text characteristics on sixth graders' inference making**

Inference generation is central to successful reading comprehension (McNamara, 2020). When reading for meaning, readers make inferences by integrating information across parts of the text, and by integrating text information with prior knowledge. The important role of inference generation in successful reading comprehension is evidenced by multiple studies demonstrating a direct relation between the two for children, adolescents, and young adults (Ahmed et al., 2016; Cain et al., 2001; Cromley et al., 2010). Good reading comprehension and inference making are essential for school-based learning in young adolescents, as the texts they encounter are increasingly complex and cover a wide range of topics. In this study, we sought to understand better the skills that influence inference making in a diverse sample of 11- to 12-year-old students in the U.S. who were either monolingual English speakers or Spanish-English bilingual speakers when in preschool. Building on research with older readers, we examined the influence of a range of reader characteristics - vocabulary and background knowledge, reading strategy awareness and use, language status, and word reading ability - on inference making from narrative and informational expository texts. This design enables us to advance our understanding of potentially malleable factors that influence inference making in a pivotal age group who are expected to apply their reading skills to learn across a variety of texts and content domains.

We contrasted inference making for narrative and informational expository text types. The two text types differ in structure and content: Narrative texts typically follow a temporal sequence, and convey information about everyday events and experiences; informational expository texts (hereafter expository) do not typically follow a timeline, provide information on an unfamiliar topic, and contain unfamiliar vocabulary (Eason et al., 2012; Graesser et al., 2004; Kulesz et al., 2016). Expository texts are important for learning across a range of content domains (Best et al., 2008; Wu et al., 2020), and a focus on applying reading to learn from text is critical for the age group in this study, young adolescent readers. Comprehension of expository texts is weaker than for narratives in general (Kulesz et al., 2016; Mar et al., 2021), and fewer inferences are made (Clinton et al., 2020). Our study provides a much-needed tightly controlled examination of the influence of text type on inference making, by comparing performance of the *same* grade 6 participants reading these different text types. Notably previous studies with this age group have either used a between group design, presenting only one text type to readers (Denton, Enos, et al., 2015) or compared reading of both genres in only a small group of students (N=9) (Lee, 2014).

The influence of inferential skills on reading comprehension increases between grades 3 to 7 (Tighe & Schatschneider, 2014). Studies with older adolescent readers demonstrate a direct relation between inference generation and reading comprehension, with inference making mediating the relation between vocabulary/background knowledge and reading comprehension (e.g., Ahmed et al., 2016). Theoretically, readers with high quality knowledge for the topic of a text can more easily integrate new information to generate inferences for comprehension and learning (McNamara & Magliano, 2009; Perfetti & Stafura, 2014). Our study was guided by models that conceptualise knowledge as a multidimensional factor (McCarthy & McNamara, 2021), and evidence of strong correlations (e.g., .59-.70) between independent measures of vocabulary and broader background knowledge (Ahmed et al., 2016; Kulesz et al., 2016). In view of these findings, we examined if participants' knowledge (vocabulary and background knowledge) modulated any effects of text type on inference making.

Vocabulary and background knowledge do not wholly explain individual differences in reading comprehension: When differences in topic-relevant knowledge are controlled, older children and better comprehenders generate more inferences from text (Barnes et al., 1996; Cain et al., 2001; Kulesz et al., 2016). Thus, we considered other potential sources of variance in inference making, in our study. Theoretical accounts of reading comprehension note the importance of strategic reading and processing capacity for the construction of a mental model of a text (McNamara & Magliano, 2009). Adolescent readers with better reading comprehension are more likely to report use of strategies that support inference making (Denton, Wolters, et al., 2015), and explicit instruction in strategy use, including those that support inference generation, results in reading comprehension gains for adolescent readers (Castells et al., 2022; Denton et al., 2014). However a direct relation between strategy use and inference making is found in some (Cromley & Azevedo, 2007), but not all (Ahmed et al., 2016) studies. Cognitive processing capacity (assessed with working memory tasks) is associated with better inference making skills and reading comprehension in sixth graders (Daugaard et al., 2017), most likely because it supports integration of information across different parts of a text (McNamara & Magliano, 2009).

Our sample included both monolingual and Spanish-English bilingual students in the U.S., the latter group entered prekindergarten (PK) speaking Spanish but subsequently experienced English-language instruction throughout schooling. There is a well-documented achievement gap between these groups (NAEP, 2020); by the end of elementary school, Spanish-English bilinguals often obtain reading comprehension scores below their age-appropriate word reading skills (Lesaux et al., 2010; Nakamoto et al., 2007). Vocabulary knowledge is one likely source of their reading comprehension difficulties (Lesaux et al., 2010; Röthlisberger et al., 2023), although not sufficient to explain fully the reading comprehension deficit of English Language Learners (Spencer & Wagner, 2017). Reading strategy knowledge or its application may also underpin poor inference making in this population (Hall et al., 2020). Our study is unique in examining the relative contributions of knowledge, and strategy awareness and use, to reading comprehension in monolingual and bilingual students.

**The current study**

We examined inference making from narrative and expository texts in U.S. monolingual English speakers and Spanish-English bilinguals in grade 6. This age group marks a critical transition between the primary and secondary grades with an increasing need to learn from text and an increasing influence on inference making on reading comprehension.

To study both accuracy and the process of inference making, each text included two pieces of information that, when integrated, supported an inference that was not explicitly stated. This information was presented in separate premise and inference-prompting sentences, and we varied task processing demands by manipulating the distance between these sentences: They were either adjacent or separated by some lines of filler text, a manipulation that affects both adult and child readers (Long & Chong, 2001; Oakhill et al., 2005). This manipulation also meant that the two critical sentences did not always occur in the same location, minimising the likelihood of readers detecting a pattern that might influence their reading behaviour. We recorded accuracy of response to a question tapping the inference for each text, the time taken to respond to the question, and sentence reading times for the premise and inference-prompting sentences.

Differences in the real-time processing of text are related to reading comprehension skill (Barnes et al., 2015; Denton, Enos, et al., 2015) suggesting that better comprehenders construct a superior mental model which facilitates integrative processing. In order to examine how individual differences in text processing were related to inference making accuracy, we constructed texts to enable a reasonably high level of accuracy under favourable processing conditions. We predicted lower levels of inference making accuracy for expository relative to narrative texts (Clinton et al., 2020) and texts with the higher processing demands (the distance manipulation). To generate the inference, readers were required to integrate information presented in separate sentences. We examined processing times for the premise and inference-prompting sentences from the same passage to index ease of integration of the inference-prompting sentence. Typically, sentences that appear earlier in a text and those that introduce new topic information are processed more slowly, as readers set up the foundation of the text or topic structure (Cirilo & Foss, 1980; Gernsbacher et al., 1990) into which other information is integrated. Further, information that more strongly prompts an inference is integrated more easily into a reader's mental model (Barnes et al., 2015). Given these conditions, we expected the inference-prompting sentence to be read more quickly than the premise, in general. The distance manipulation (near vs far) served to indicate if cognitive processing demands (greater in the far condition) influenced this process.

A critical focus was to understand better the source of variation in accuracy and processing times. To do this, we examined the influence of reading and reading-related skills both theoretically and empirically related to inference making in previous work: vocabulary and background knowledge and awareness and use of reading strategies. Given the literature reviewed, we included multiple measures of vocabulary and background knowledge and used factor analysis to inform their treatment in our analyses. We also examined the influence of sight word reading on inference task performance. Although not a focus of instruction in grade six, word reading continues to predict reading comprehension and inference making in adolescence (Ahmed et al., 2016; Barnes et al., 2015), making it a potential explanatory factor in our reading-based inference paradigm. We expected vocabulary and background knowledge, and reading strategies, to relate more strongly to performance on the expository texts (Denton, Enos, et al., 2015) and lower levels of inference accuracy, in general, in the bilingual sample (Lesaux et al., 2010).

# Method

## Participants

Eighty-six monolingual English speakers and 95 children who entered pre-kindergarten (PK) as Spanish-English bilingual speakers participated in this study when they were in grade six (Spring 2018). Our analyses included only those with full data resulting in a sample of 152 participants. See Table 1 for sample demographics. The study conforms to the US Federal Policy for the Protection of Human Subjects and was approved by the Institutional Review Board or Research Ethics Committee at each university. Informed consent was provided by legal guardians and children gave their assent prior to participation.

## Measures and Procedure

Children were tested individually in a quiet room in their school or at a university lab. See Table 2 for scores and item reliabilities of all measures.

### *Sight word reading.*

Children completed the Sight Word Efficiency subtest of the Test of Word Reading Efficiency – Second Edition (TOWRE-2; Torgesen et al., 2012), which measures the number of English words, ranging from high to low frequency, pronounced correctly in 45 seconds. The test was administered and scored in line with the manual.

***Vocabulary.***

Three measures of vocabulary were administered. The Peabody Picture Vocabulary-4 (PPVT-4; Dunn & Dunn, 2007) assessed understanding of single spoken words. The Expressive Vocabulary Test-2 (EVT-2; Williams, 2007) assessed single-word productive vocabulary. The CELF-4 (Semel & Wiig, 2006) Word Definitions Subtest assessed depth of vocabulary knowledge.

***Background knowledge.***

Participants completed 32 open-ended questions assessing knowledge of topics and key concepts in science, social science, and biography (e.g., "Why does your body need oxygen?" "What is an archaeologist?"). These questions were selected from those relating to passages in the Qualitative Reading Inventory (Leslie & Caldwell, 2011) that formed a measure in the larger test battery. Each response was scored from 0-3 using a rubric developed from the responses of 28 children, with high inter-rater reliability ($r$ = .87).

***Reading strategy awareness and use.***

Participants completed a 49-item questionnaire, adapted from Denton, Wolters, et al. (2015), to assess awareness and use of reading strategies (e.g., "I take notes while I am reading"), for different reading scenarios, such as reading a section of a social studies textbook to prepare for a class discussion. For each item, participants responded yes/no to indicate they would use that strategy in that scenario. Responses to the 14 items that contributed to the "Evaluation and Integration" subscale were summed (1 for yes responses, 0 for no responses) and used in the analysis.

***Inference*.**

Each child read 12 seven-sentence narrative texts and 12 seven-sentence expository texts written for this age group (examples in Table S.1 Supplementary Materials). The narrative texts concerned human characters and focussed on typical activities for children, such as parties, family, and friendships (some adapted from Pike et al. (2010). The expository texts had a science theme and each focussed on one topic such as plants, minerals or metals. Coh-metrix 3.0 text analysis (Graesser et al., 2011) confirmed no differences between sets for word concreteness ($M_{narrative}$ = 93.71 (SD=14.81); $M_{expository}$ = 84.96 (SD=22.66); $t(46) = 1.57$, $p = .12$, but higher narrativity scores for narratives ($M_{narrative}$ = 63.50 (SD=20.14); $M_{expository}$ = 19.54 (SD=13.35); $t(46) = 8.91$, $p < .001$.

Children read each text, sentence by sentence, on a laptop via E-Prime 3.0 (Psychology Software Tools, 2016) and advanced to each new sentence by pressing a key on an E-Prime button box. Participants could not return to a previous sentence. The reading time for each sentence was recorded. Each text required an inference to integrate the premise and inference-prompting sentences (hereafter referred to as critical sentences). The processing load was manipulated by placing the two sentences required to generate the inference either adjacent to each other (near condition) or 1-3 sentences apart (far condition), counterbalanced across two presentation lists so that children read each text in only one condition. There were no differences by text type in total words per text ($M_{narrative}$ = 83.42 (SD=10.17), $M_{expository}$ = 88.33 (SD=7.01), $t(22) = 1.38$, $p = .18$), nor intervening words in the far condition ($M_{narrative}$ = 30.75 (SD=9.80), $M_{expository}$ = 30.83 (SD=12.13), $t(22) = .02$, $p = .99$).

After each text, participants responded yes/no to a question that connected the two critical sentences (see Singer & Halldorson (1996) for a similar test of inference validation). There was no difference by text type in the frequency of words in the questions ($M_{narrative}$ = 2.81; $M_{expository}$ = 2.76; Graesser et al., 2011). The time taken to read and respond to the question was recorded. For each text type, six texts had questions requiring a yes response, and six a no response. Two practice texts, with feedback, were first completed. Accuracy for the question, time taken to respond to the question, and time to read the two sentences critical to generating the inference were used as outcome variables in separate analyses.

**Overview of Data Analysis**

Inference question accuracy, inference question reading and response times, and critical sentence reading times were analysed with (Generalised) Linear Mixed-effects Models (GLMMs) using the lme4 package for R (Bates et al., 2015). Models were fitted to estimate the predicted (fixed) effects of key variables (word reading ability, knowledge, reading strategies, text type, and distance condition) and their interactions while taking into account random effects associated with differences between sampled children or texts. Categorical fixed effects were contrast coded to interpret the lower order (main) effects: Text type (Narrative = +1, Expository = -1); Distance condition (Near = +1, Far = -1) and, for the sentence reading times analysis only, Sentence type (Inference = +1, Premise = -1). Word reading, knowledge, and reading strategy scores were centered and scaled. Models specified with maximal random effects structure (Barr et al., 2013) did not converge, so we report models which did converge by reducing model complexity (Brauer & Curtin, 2018; Matuschek et al., 2017).

## Results

One bilingual participant and two monolingual participants did not complete the inference expository task; their data for the inference narrative task was retained. Our original intention (see pre-registration: https://osf.io/j7zps) was to contrast U.S. Spanish-English bilingual and monolingual readers to examine the contribution of language status (a binary variable) to inference generation alongside other potentially *malleable* reader characteristics - knowledge and reading strategies - that could be targeted in interventions.

We adapted our approach in the light of sizeable heterogeneity on the reader characteristics in each of the two groups and overlapping distributions (see Table 2). Therefore, we examined how variation in reader characteristics across the whole sample were driving individual differences in inference making, in addition to contrasting language groups. Denton and colleagues (Denton, Enos et al., 2015) adopt a similar approach including, but not analytically contrasting, U.S. monolingual and bilingual students.

**Individual Difference Measures**

Monolingual participants obtained significantly higher (raw) scores on all vocabulary measures and the background knowledge measure, but the groups did not differ significantly for word reading and reading strategy awareness and use. Correlations between the vocabulary and the background knowledge measures were all > 0.7. An exploratory factor analysis revealed a single factor (Eigenvalue > 1.0; task loadings > 0.70; proportion of variance explained = 79.20). Taking this as evidence for a single construct, a knowledge factor score comprising these four measures was used in all subsequent analyses.

The following analyses assess how performance on the assessments of word reading, knowledge, and reading strategies relate to performance on the inference task. In our pre-registration (https://osf.io/j7zps), we planned to include language status as a fixed effect in our models. Therefore, for each inference outcome we first conducted the modelling with language status included as a predictor. As noted above, there were large and significant differences in performance between the language status groups on the individual measures that contributed to the knowledge factor score. The inclusion of language status as a fixed effect in all models did not improve model fit and, when both language status and knowledge were included in the models simultaneously, the influence of both predictors was suppressed. Below we report models excluding language status to examine the effect of (malleable) individual differences on the inference task below, and report model summaries with language status in our Supplementary Materials. The same pattern of effects was evident in models including and excluding language status.

**Responses to Inference Questions**

***Accuracy***

The mean proportions of correct responses to the inference question are reported in Table 3 and model summary in Table 4. The first column of Table 4 provides the coefficient estimates for the effects (*B*). The significant positive value for the intercept (coefficient B = 2.00) shows an overall 0.88[1] probability correct. The key predictor was knowledge: participants with higher knowledge scores were more likely to obtain higher inference scores than those with lower knowledge scores. As depicted in Figure 1, there was a steeper decline in predicted performance for individuals scoring below average for knowledge. Text type had a significant effect on question accuracy: narrative scores were 4% more accurate than the grand mean, and expository 4% less accurate (Table 3). There were no significant interactions.

***Inference Question Reading Time***

Reading times (milliseconds per word) to questions answered correctly (82% of all items) were analysed using linear mixed-effects models (means in Table 5; model summary

[1] The 2.00 intercept coefficient represents odds of exp(2.00) = 7.39 which in turn represents an overall probability of being correct of *p* = 7.39/ 1 + 7.39 = .88 which is our models' best estimate for the grand mean for the data (and is in line with our sample grand mean of ( .86 + .77)/2 = .82 (see Table 3)). Using our model coefficients, the log odds of being correct for narrative is 2.00 + 0.40 = 2.40 and for expository is 2.00 – 0.40 = 1.6. The odds of being correct for each text type are narrative exp(2.40) = 11.02 and expository exp(1.6) = 4.95 The probability correct for narrative is 11.02/(1 + 11.02) = .92 and expository is 4.95/(1 + 4.95) = .83. Therefore the .40 beta coefficient represents an estimated 4% change in accuracy either side of the grand mean for the two text types. The 4% change in accuracy is reflected in differences between the actual sample means narrative = .86 and expository = .77 around the actual sample grand mean .82.

Table 6). The model included the same fixed effects, interactions, and random intercepts as the question accuracy model but the random slopes differed; there were random effects terms to account for between-items differences in the effect of knowledge and between-participant differences in the effect of text type.

There was a significant effect of knowledge; participants with higher knowledge scores read and responded to questions more quickly than those with lower knowledge scores. There was a significant interaction between text type and word reading. Examination of the interaction plot (Figure 2) and analysis of the data after sub-setting by text type indicated that better word reading was associated with faster times for questions following narratives ($B$ = -66.80 ($SE$ = 28.54), $t$ = 2.34), but there was no association between word reading and question response times for expository texts ($B$ = -5.81 ($S E$= 31.33), $t$ = -0.19). The key predictor of the time taken to read and respond to the questions following expository texts was knowledge ($B$ = -141.93 ($SE$ = 38.59), $t$ = -3.68. See Supplementary Materials Tables S.5 and S.6 for model summaries.

**Sentence Reading Times**

Reading times for the premise sentence and the inference-prompting sentence were converted to milliseconds per word and analysed using linear mixed-effects models (see Table 7 for means; Table 8 for model summary). To obtain a within-text comparison, we compared the reading time for the premise and inference-prompting sentences from the same passage. As above, the models included the same fixed effects, interactions, and random intercepts, with the addition of sentence type as a fixed effect. There was one random effect included to account for between-participant differences in the slope of the effect of text type.

There were main effects of word reading and knowledge because readers with better word reading and/or knowledge read the sentences more quickly in general. There was a main effect of sentence type; the inference-prompting sentence was read more quickly than the premise sentence. There was a significant three-way interaction between knowledge, distance condition, and sentence type, which is plotted in Figure 3. There was a stronger effect of knowledge on the inference-prompting sentence reading time in the near than in the far condition; participants with knowledge scores more than one standard deviation below the mean did not show a facilitation effect in the near condition, such that both sentences were read at a similar rate. In general, narrative texts were read more quickly than expository texts, but the effect of genre did not reach statistical significance.

## Discussion

This study provides new information about the skills that predict inference making for different text types in 11- to -12-year-olds, an age group where inference making is increasingly important due to advanced school learning demands. Participants were more likely to correctly answer an inference question after reading narrative than expository texts. Greater knowledge was associated with increased probability of answering an inference question correctly, faster response times to correct questions, and faster reading of critical information in the text, but participants' awareness and use of reading strategies did not influence performance on any of our measures. Of note, models predicting inference performance with either knowledge or participant language-status (monolingual, bilingual) showed the same pattern of results.

We reproduced the advantage for accuracy of inference making from narrative texts (Clinton et al., 2020) and the strong and positive effects of reader knowledge on inference making (Barnes et al., 2015; Kulesz et al., 2016). In contrast to Best et al.'s (2008) study of (slightly) younger children, we did not find an interaction between readers' knowledge and text type on either question answering accuracy or speed. Knowledge is hypothesised to play a greater role in expository text comprehension because such passages typically convey unfamiliar content (Graesser et al., 2004). Like Best et al. (2008), our measure of knowledge was not specific to passage content, so differences in the measurement of knowledge do not provide a satisfactory explanation for differences between studies.

A key difference between our study and Best et al. (2008) is that our two text types did not differ greatly in difficulty: The difference in overall percent correct between our narrative and expository texts was just 9% (86% vs 77%) compared to 23% (72% vs 49%) in Best et al. (2008). Our text types did not differ in word concreteness (see method), whereas a feature of expository texts is often a higher frequency of abstract concepts compared with narratives (Graesser et al., 2004). Thus, the differences between our text types may not reflect those found in materials used in other studies, limiting direct comparisons and, critically, a more precise understanding of the factors that influence the reported greater difficulty of informational texts. Related research demonstrates genre differences that are independent of text features such as sentence length, word frequency, and cohesion (Kulesz et al., 2016). Future work needs to build on these collected findings to disentangle the effects of a range of text features, in addition to text type, and the influence of reader characteristics, such as knowledge, to understand how these influence understanding of different text genres (Kulesz et al., 2016; Ozuru et al., 2009). Knowledge has been found to exert a greater influence on inference making knowledge under certain conditions (Kulesz et al., 2016; Ozuru et al., 2009), which prompts the need to discover critical reader x text interactive effects.

Our sample included two groups who differed in language status at pre-kindergarten entry. In line with a growing evidence base, the bilingual sample had age appropriate word reading scores but, on average, lower vocabulary and background knowledge scores than monolinguals (Lesaux et al., 2010). Overall, the bilingual sample had lower accuracy on the inference task and took longer to answer questions. We chose to report the models with knowledge, rather than language status, as a predictor to highlight a potentially important malleable factor that could be targeted in instruction. Models predicting inference performance with either knowledge or participant language-status yielded the same pattern of results. These findings indicate weaker knowledge as a potential source of the bilingual students' poorer performance, in line with other research showing the predictive power of vocabulary and background knowledge on text comprehension (Kulesz et al., 2016).

Our study was not designed to test for causal relations between participant variables and inference making, but the proposal for a causal relation between knowledge and inference making is supported by a strong body of correlational research (McCarthy & McNamara, 2021), as well as longitudinal studies indicating reciprocal relations between vocabulary knowledge and inference making in younger children (Language and Reading Research Consortium (LARRC) et al., 2019). Meta-analytic reviews demonstrate a positive effect of vocabulary instruction on the comprehension of texts containing those words, indicating the importance of text-specific knowledge (Wright & Cervetti, 2017), and positive impacts of academic vocabulary instruction for readers whose home language differs from the language of schooling (Lesaux et al., 2014). Furthermore, when students' ability to recall an explicitly-taught knowledge base relating to story-specific content predicts their ability to make inferences using that knowledge base (Barth et al., 2021). This body of research indicates substantial potential for instruction in knowledge and calls for intervention studies that test the impact of (different methods of) knowledge instruction on inference making to test the potential causal relation between knowledge and inference making.

We examined the influence of working memory (or processing load) by manipulating the distance between the two sentences containing information critical to the inference. We did not find an overall effect of distance on sentence reading times, but this manipulation influenced sentence reading times in combination with knowledge and sentence type. When the information to be integrated was presented in sentences separated by filler text (far condition), the influence readers' knowledge was comparable for the two sentences. When the information to be integrated was presented in adjacent sentences (near condition), readers with lower levels of conceptual knowledge did not show a processing advantage for the inference-prompting information. This latter finding can be related to memory-based accounts of text processing that propose inference making is facilitated when the relevant textual information and semantic and general knowledge are activated and available in working memory (see van den Broek et al. (2005) for a review).

**Limitations and future research**

In addition to the limitations discussed above, we highlight methodological limitations that should be addressed in future studies. Awareness and reported use of reading strategies did not explain variation in performance on our inference task (see also Ahmed et al., 2016). Given the high performance accuracy, our study may be limited in the opportunity for readers to apply strategic processes. In addition, we used a paper-and-pen assessment of reading strategies but a timed measure of inference making. The assessment of strategy use in real-time processing of text, such as a think-aloud task, may be a more appropriate and sensitive measure (Denton, Enos, et al., 2015), alongside longer and more challenging texts. Similarly, our measurement of inference making through yes/no question responses was influenced by the need for a timed measure; untimed measures, such as open-ended questions or retells, may provide additional indices of the use of knowledge in inference making.

**Implications and conclusions**

There are educational implications that stem from this work. Despite the importance given to knowledge in models of reading comprehension (McCarthy & McNamara, 2021; Perfetti & Stafura, 2014), it has been noted that teaching to support the acquisition of knowledge across sources and topics has not been a primary focus of language and reading curricula (Elleman & Compton, 2017). Our findings support the calls for a focus on the acquisition and use of knowledge, alongside other literacy skills, in reading comprehension instruction (Elleman et al., 2009; Hwang et al., 2022). Although strategic reading was not a unique prediction of variance in our study, we note that reading in general, and strategic reading in particular, may support vocabulary and knowledge gains (Cain et al., 2004; McCarthy & McNamara, 2021). Finally, given the weaker reading comprehension skills reported in the Spanish-English bilingual readers in this age group, together with their weaker knowledge, we propose that building knowledge is a focus for instruction with this group (Barth et al., 2021).

We have established that knowledge is a significant predictor of inference making in 11- to 12-year-olds explaining performance on both accuracy and processing measures. Knowledge is both a product and predictor of reading comprehension, and inference making itself, in younger age groups (Cain & Oakhill, 2012; LARRC et al., 2019). This study provides further evidence to support the call for instruction to support acquisition and use of knowledge in the service of reading comprehension.

**References**

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*, 66-82. https://doi.org/10.1016/j.cedpsych.2016.02.002

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., & 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

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barth, A. E., Daniel, J., Roberts, G., Vaughn, S., Barnes, M. A., Ankrum, E., & Kincaid, H. (2021). The role of knowledge availability in forming inferences with rural middle grade English learners. *Learning and Individual Differences*, *88*, 102006. https://doi.org/10.1016/j.lindif.2021.102006

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

Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, *29*(2), 137-164. https://doi.org/10.1080/02702710801963951

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389-411. https://doi.org/10.1037/met0000159

Cain, K., Lemmon, K., & Oakhill, J. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, *96*(4), 671-681. https://doi.org/10.1037/0022-0663.96.4.671

Cain, K., & Oakhill, J. (2012). Reading comprehension development from seven to fourteen years: Implications for assessment. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring Up: Advances in How to Assess Reading Ability* (pp. 59--76). Rowman & Littlefield Education.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference making ability and their relation to knowledge. *Memory and Cognition*, *29*(6), 850-859. https://doi.org/10.3758/BF03196414

Castells, N., Minguela, M., Solé, I., Miras, M., Nadal, E., & Rijlaarsdam, G. (2022). Improving questioning–answering strategies in learning from multiple complementary texts: An intervention study. *Reading Research Quarterly*, *57*(3), 879-912. https://doi.org/10.1007/s11145-020-10044-2

Cirilo, R. K., & Foss, D. J. (1980). Text structure and reading time for sentences. *Journal of Verbal Learning and Verbal Behaviour*, *19*, 96-109. https://doi.org/10.1016/S0022-5371(80)90560-5

Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E., & Seipel, B. (2020). Inferential comprehension differences between narrative and expository texts: a systematic review and meta-analysis. *Reading and Writing*, *33*, 2223-2248. https://doi.org/10.1007/s11145-020-10044-2

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*(2), 311–325. https://doi.org/10.1037/0022-0663.99.2.311

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *102*(3), 687–700. https://doi.org/10.1037/a0019452

Daugaard, H., Cain, K., & Elbro, C. (2017). From words to text: Inference making mediates the role of vocabulary in children's reading comprehension. *Reading and Writing*, *30*, 1773-1788. https://doi.org/10.1007/s11145-017-9752-2

Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., Fletcher, J. F., & Carter, S. (2015). Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly*, *50*(4), 393-416. https://doi.org/10.1002/rrq.105

Denton, C. A., Fletcher, J. M., Taylor, W. P., Barth, A. E., & Vaughn, S. (2014). An experimental evaluation of guided reading and explicit interventions for primary-grade students at-risk for reading difficulties. *Journal of Research on Educational Effectiveness*, *7*(3), 268-293. https://doi.org/10.1080/19345747.2014.906010

Denton, C. A., Wolters, C. A., York, M. J., Swanson, E., Kulesz, P. A., & Francis, D. J. (2015). Adolescents' use of reading comprehension strategies: Differences related to reading proficiency, grade level, and gender. *Learning and Individual Differences*, *37*, 81-95. https://doi.org/10.1016/j.lindif.2014.11.016

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test--Fourth Edition*. Pearson Assessment.

Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader–text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, *104*(3), 515-528. https://doi.org/10.1037/a0027182

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

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44. https://doi.org/10.1080/19345740802539200

Gernsbacher, M. A., Varner, K. R., & Faust, M. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 430-445. https://doi.org/10.1037/0278-7393.16.3.430

Graesser, A., McNamara, D. S., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers*, *36*, 193-202. https://doi.org/10.3758/BF03195564

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, *40*(5), 223-234. https://doi.org/10.3102/0013189X11413260

Hall, C., Vaughn, S., Barnes, M. A., Stewart, A. A., Austin, C. R., & Roberts, G. (2020). The effects of inference instruction on the reading comprehension of English learners with reading comprehension difficulties. *Remedial and special education*, *41*, 259–270. https://doi.org/10.1177/0741932518824983

Hwang, H. J., Cabell, S. Q., & Joyner, R. E. (2022). Effects of integrated literacy and content-area instruction on vocabulary and comprehension in the elementary years: A meta-analysis. *Scientific Studies of Reading*, *26*(3), 223-249. https://doi.org/10.1080/10888438.2021.1954005

Kulesz, P. A., Francis, D. J., Barnes, M. A., & Fletcher, J. M. (2016). The influence of properties of the test and their interactions with reader characteristics on reading comprehension: An explanatory item response study. *Journal of Educational Psychology*, *108*(8), 1078–1097. https://doi.org/10.1037/edu0000126

Language and Reading Research Consortium (LARRC), Muijselaar, M. M. L., & Currie, N. K. (2019). Inference making in young children: the concurrent and longitudinal contributions of verbal working memory and vocabulary. *Journal of Educational Psychology*, *111*(8), 1416-1431. https://doi.org/10.1037/edu0000342

Lee, H. C. (2014). Inferencing behaviour of ESL readers. *Electronic Journal of Foreign Language Teaching*, *11*(1), 21–37.

Lesaux, N. K., Crosson, A. C., Kieffer, M. J., & Pierce, M. (2010). Uneven profiles: Language minority learners' word reading, vocabulary, and reading comprehension skills. *Journal of Applied Developmental Psychology*, *31*(6), 475-483. https://doi.org/10.1016/j.appdev.2010.09.004

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, *51*(6), 1159-1194. https://doi.org/10.3102/0002831214532165

Leslie, L., & Caldwell, J. S. (2011). *Qualitative reading inventory* (5 ed.). Pearson.

Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1424-1429. https://doi.org/10.1037/0278-7393.27.6.1424

Mar, R. A., Li, J., Nguyen, A. T., & Ta, C. P. (2021). Memory and comprehension of narrative versus expository texts: A meta-analysis. *Psychonomic Bulletin and Review*, *28*, 732-749. https://doi.org/10.3758/s13423-020-01853-1

Matuschek, H., Kliegl, R., Vasishth, S., & Baayen, H. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *95*, 305-315. https://doi.org/10.1016/j.jml.2017.01.001

McCarthy, K. S., & McNamara, D. S. (2021). The multidimensional knowledge in text comprehension framework. *Educational Psychologist*, *56*(3), 196-214. https://doi.org/10.1080/00461520.2021.1872379

McNamara, D. S. (2020). If integration Is the keystone of comprehension: Inferencing Is the key. *Discourse Processes*, *58*, 86-91. https://doi.org/10.1080/0163853X.2020.1788323

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 51, pp. 297-384). Academic Press.

NAEP (2020). *Nation's report card. National Assessment of Educational Progress.* Retrieved 04/08/2022 from https://www.nationsreportcard.gov/ltt/reading/student-group-scores/?age=9

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing*, *20*, 691-719. https://doi.org/10.1007/s11145-006-9045-7

Oakhill, J. V., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing*, *18*, 657-713. https://doi.org/10.1007/s11145-005-3355-z

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, *19*(3), 228e242. https://doi.org/10.1016/j.learninstruc.2008.04.003

Perfetti, C. A., & Stafura, J. (2014). Reading comprehension: Including word knowledge in a theoretical framework. *Scientific Studies of Reading*, *18*(1), 22-37. https://doi.org/10.1080/10888438.2013.827687

Pike, M. M., Barnes, M. A., & Barron, R. W. (2010). The role of illustrations in children's inferential comprehension. *Journal of Experimental Child Psychology*, *105*(3), 243-255. https://doi.org/10.1016/j.jecp.2009.10.006

Psychology Software Tools, I. (2016). *E-Prime 3.0*. In https://www.pstnet.com

Röthlisberger, M., Zangger, C., & Juska-Bacher, J. (2023). The role of vocabulary components in second language learners early reading. *Journal of Research in Reading*, *46*(1), 1-21. https://doi.org/https://doi.org/10.1111/1467-9817.12411

Semel, E., & Wiig, E. (2006). *Clinical Evaluation of Language Fundamentals: 4*. Pearson Assessment.

Singer, M., & Halldorson, M. (1996). Constructing and validating motive bridging inferences. *Cognitive Psychology*, *30*(1), 1-38. https://doi.org/10.1006/cogp.1996.0001

Spencer, M., & Wagner, R. K. (2017). The comprehension problems for second-language learners with poor reading comprehension despite adequate decoding: A meta-analysis. *Journal of Research in Reading*, *40*(2), 199-217. https://doi.org/10.1111/1467-9817.12080

Tighe, E. L., & Schatschneider, C. (2014). A dominance analysis approach to determining predictor importance in third, seventh, and tenth grade reading comprehension skills. *Reading and Writing*, *27*, 101-127. https://doi.org/10.1007/s11145-013-9435-6

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of Word Reading Efficiency (TOWRE).* (2 ed.). Pro-Ed.

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist approaches in accounts of reading comprehension. *Discourse Processes*, *39*, 299–316. https://doi.org/10.1080/0163853X.2005.9651685

Williams, K. T. (2007). *Expressive Vocabulary Test (2nd ed.)*. Pearson.

Wright, T. S., & Cervetti, G. N. (2017). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, *52*(2), 203-226. https://doi.org/10.1002/rrq.163

Wu, Y., Barquero, L. A., Pickren, S. E., Barber, A. T., & Cutting, L. E. (2020). The relationship between cognitive skills and reading comprehension of narrative and expository texts: A longitudinal study from Grade 1 to Grade 4. *Learning and Individual Differences*, *80*, 101848. https://doi.org/10.1016/j.lindif.2020.101848

**Table 1**

*Demographic Characteristics of English Monolingual and Spanish-English Bilingual Participants*

| | | English Monolingual | Spanish-English Bilingual |
|---|---|---|---|
| N (%female) | | 71 (54%) | 81 (35%) |
| Age (years; months) | | 12;1 | 12;1 |
| U.S. state of residence | | Arizona, Kansas, Nebraska | Arizona |
| Income | | | |
| | < 20k | 0 | 32 |
| | 20,001-40k | 8 | 37 |
| | 40,001-60k | 8 | 7 |
| | 60,001-80k | 9 | 4 |
| | >80k | 46 | 1 |
| Mother/Female Guardian Education Level[a] | | | |
| | <High school | 0 | 53 |
| | High school | 3 | 15 |
| | Some college | 13 | 3 |
| | Associates/Technical degree | 6 | 4 |
| | Bachelor's degree | 23 | 4 |
| | Post graduate degree | 26 | 1 |
| Race N(%) | | | |

| | | |
|---|---|---|
| White | 65 (91.5%) | 55 (67.9%) |
| Multiracial | 2 (2.8%) | 16 (19.8%) |
| Other/unknown | 4 (5.6%) | 7 (8.6%) |
| Not reported | 0 (0%) | 3 (3.7%) |
| Ethnicity N(%) | | |
| Hispanic or Latino | 2 | 46 |
| Not Hispanic or Latino | 8 | 18 |
| Not reported | 9 | 2 |
| Free/reduced lunch | 9 | 72 |

[a]Missing data for one participant.

**Table 2**

*Performance on Individual Difference Measures, Reliability, and Effect Sizes for Language Status Group Comparisons*

| | Language Status | | Reliability (*r*) | Effect size (η2) |
|---|---|---|---|---|
| | Monolingual | Bilingual | | |
| N | 71 | 81 | | |
| Word reading | 78.94 (10.78) | 75.78 (8.92) | .93 | .03 |
| (TOWRE SWE) | 45-106 | 50-94 | | |
| Standardised score | 105.06 (15.24) | 99.67 (12.90) | | |
| | 68 -143 | 73 -130 | | |
| Reading strategy | 5.54 (3.58) | 6.63 (3.95) | | .02 |
| knowledge | 1-14 | 0-14 | | |
| [a]Receptive vocabulary | 185.44 (13.93) | 158.30 (19.06) | [b].95 | .40*** |
| (PPVT) | 145-207 | 112-204 | | |
| Standardised score | 113.85 (11.48) | 94.04 (12.55) | | |
| | 84 -133 | 65 -128 | | |
| [a]Expressive vocabulary | 139.00 (12.79) | 116.51 (15.28) | .94 - .97 | .39*** |
| (EVT) | 112-162 | 78-151 | | |
| Standardised score | 110.89 (12.02) | 92.31 (11.83) | | |
| | 88 -133 | 65 -120 | | |
| [a]Word definitions | 29.34 (6.36) | 16.80 (7.59) | .83 | .44*** |
| (CELF) | 16-44 | 5-34 | | |

| | | | |
|---|---|---|---|
| [a]Background knowledge (QRI, proportionate score) | 0.44 (0.13) .09-.71 | .25 (0.11) .04-.56 | .36*** |
| Knowledge (factor score) | 0.71 (0.62) -.67-1.89 | -0.62 (0.78) -2.21-1.44 | .47*** |

*Note*. N = 152. Raw scores, standard deviation and range are reported for each measure, and standardised scores where available. For one-way ANOVAs comparing the two groups on the raw scores of the individual difference measures, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Test-retest from manual unless stated otherwise. TOWRE SWE = TOWRE-2 Sight Word Efficiency subtest (Torgesen, Wagner, & Rashotte, 1999); EVT = Expressive Vocabulary Test-2 (EVT-2; Williams, 2007); CELF = CELF-4 (Semel & Wiig, 2006); PPVT = PPVT-4 (Dunn & Dunn, 2007); QRI = Qualitative Reading Inventory (Leslie & Caldwell, 2011). [a] Measures contributing to the Knowledge factor score. [b] Split-half reliability.

**Table 3**

*Mean Proportion of Correct Responses (and Standard Deviations) for the Inference Question*

| Text Type | Distance Condition | Language Status | | Total | |
|---|---|---|---|---|---|
| | | Monolingual | Bilingual | | |
| Narrative | Near | 0.90 (0.30) | 0.85 (0.36) | 0.87 (0.34) | 0.86 (0.35) |
| | Far | 0.88 (0.33) | 0.81 (0.39) | 0.84 (0.37) | |
| Expository | Near | 0.86 (0.35) | 0.70 (0.46) | 0.77 (0.42) | 0.77 (0.42) |
| | Far | 0.85(0.35) | 0.68 (0.47) | 0.76 (0.43) | |
| Total | | 0.87 (0.34) | 0.76 (0.43) | | |

*Note*. Means for each language status group are included for comparison.

**Table 4**

*Summary GLMM for (log odds) Inference Question Accuracy Without Language Status as a Fixed Effect.*

| Fixed effects | Estimated coefficient | *SE* | *z* | *p* |
|---|---|---|---|---|
| (Intercept)[2] | 2.00 | 0.20 | 9.87 | <.001 |
| Reading strategy | -0.02 | 0.08 | -0.27 | .79 |
| Word reading | 0.03 | 0.08 | 0.35 | .73 |
| **Knowledge** | **0.79** | **0.10** | **7.89** | **<.001** |
| **Text type** | **0.40** | **0.20** | **2.05** | **.04** |
| Distance condition | 0.08 | 0.05 | 1.65 | .10 |
| Reading strategy x Text type | 0.05 | 0.06 | 0.88 | .38 |
| Word reading x Text type | -0.02 | 0.06 | -0.35 | .72 |
| Knowledge x Text type | -0.08 | 0.08 | -1.03 | .30 |
| Reading strategy x Distance condition | 0.01 | 0.05 | 0.21 | .84 |
| Word reading x Distance condition | -0.09 | 0.06 | -1.71 | .09 |
| Knowledge x Distance condition | -0.01 | 0.06 | -0.20 | .84 |
| Text type x Distance condition | | | | |
| Reading strategy x Text type x Distance | 0.06 | 0.05 | 1.25 | .21 |
| condition | 0.03 | 0.05 | 0.66 | .51 |
| Word reading x Text type x Distance | | | | |
| condition | -0.04 | 0.05 | -0.70 | .48 |

| | | | | | |
|---|---|---|---|---|---|
| Knowledge x Text type x Distance condition | | 0.09 | 0.05 | 1.57 | .12 |
| Random effects | | | | Variance | *SD* |
| | Participant | (intercept) | | 0.41 | 0.64 |
| | | Text type | | 0.08 | 0.29 |
| | Item | (intercept) | | 0.81 | 0.90 |
| | | Reading strategy | | 0.02 | 0.13 |
| | | Knowledge | | 0.05 | 0.23 |
| $R^2$ marginal[b] = 0.15, $R^2$ conditional[c] = 0.40 | | | | | |

*Note.* Observations = 3612; Participants = 152; Items = 24; $R^2$ calculated using the MuMIn package in R; Effects in bold are statistically significant. [a] Three participants had missing expository data (one bilingual). [b] Variance explained by the fixed effects. [c] Variance explained by the entire model including both fixed and random effects.

**Table 5**

*Inference Question Reading Times for Correctly Answered Items (Milliseconds per Word)*

| Text type | Distance Condition | Language Status | | Total | |
|---|---|---|---|---|---|
| | | Monolingual | Bilingual | | |
| Narrative | Near | 522.44 | 697.68 | 613.27 | 640.97 |
| | | (368.98) | (567.64) | (489.87) | (565.28) |
| | Far | 589.01 | 745.94 | 669.62 | |
| | | (667.19) | (589.47) | (632.94) | |
| Expository | Near | 593.94 | 840.05 | 714.15 | 718.76 |
| | | (446.91) | (1014.87) | (787.08) | (718.86) |
| | Far | 630.82 | 824.42 | 723.47 | |
| | | (489.67) | (763.11) | (642.24) | |
| Total | | 582.91 | 771.44 | | |
| | | (506.39) | (742.79) | | |

*Note*. Means for each language status group are included for comparison.

**f**

*Summary LMM for Inference Question Reading Time (Milliseconds per Word) Without Language Status as a Fixed Effect*

| Fixed effects | Estimated coefficient | *SE* | *t* | *p* |
|---|---|---|---|---|
| (Intercept) | 724.85 | 70.89 | 10.23 | <.001 |
| Reading strategy | -11.97 | 23.34 | -0.51 | .61 |
| Word reading | -38.10 | 26.28 | -1.45 | .15 |
| **Knowledge** | **-106.71** | **31.12** | **-3.32** | **.001** |
| Text type | -66.41 | 68.07 | -0.98 | .34 |
| Distance condition (Near) | -13.65 | 9.26 | -1.47 | .14 |
| Reading strategy x Text type | 7.05 | 12.66 | 0.56 | .58 |
| **Word reading x Text type** | **-28.96** | **14.00** | **-2.07** | **.04** |
| Knowledge x Text type | 32.43 | 23.16 | 1.40 | .17 |
| Reading strategy x Distance condition | -9.02 | 9.36 | -0.96 | .34 |
| Word reading x Distance condition | 9.74 | 10.47 | 0.93 | .35 |
| Knowledge x Distance condition | -13.38 | 10.58 | -1.27 | .21 |
| Text type x Distance condition | -5.79 | 9.27 | -0.62 | .53 |
| Reading strategy x Text type x Distance condition | 0.68 | 9.36 | 0.07 | .94 |
| Word reading x Text type x Distance | 2.41 | 10.48 | 0.23 | .82 |
| condition | 0.75 | 10.60 | 0.07 | .94 |
| Knowledge x Text type x Distance condition | | | | |

| Random effects | | Variance | *SD* |
|---|---|---|---|
| Participant | (intercept) | 69713 | 264.03 |
| | Text type | 10195 | 100.97 |
| Item | (intercept) | 107384 | 327.70 |
| | Knowledge | 8036 | 89.64 |
| $R^2$ marginal[b] = 0.05, $R^2$ conditional[c] = 0.47 | | | |

*Note.* Observations = 2929; Participants = 152; Items = 24; $R^2$ calculated using the MuMIn package in R; Effects in bold are statistically significant. [a] Correct responses only; three participants had missing expository data (one bilingual participant). [b] Variance explained by the fixed effects. [c] Variance explained by the entire model including both fixed and random effects.

**Table 7**

*Premise Sentence and Inference Sentence Reading Times: Correct Responses (Milliseconds per Word)*

| Text type | Distance condition | Language Status | | | Totals | |
|---|---|---|---|---|---|---|
| | | Monolingual | Bilingual | | | |
| Narrative | Near | | | | | |
| | Premise | 324.62 (181.69) | 405.82 (203.42) | 366.70 (197.36) | 358.27 (205.72) | 362.90 (201.62) |
| | Inference | 301.61 (158.86) | 394.64 (245.89) | 349.83 (213.55) | | |
| | Far | | | | | |
| | Premise | 335.75 (186.63) | 428.92 (208.41) | 383.61 (203.40) | 367.68 (197.24) | |
| | Inference | 312.96 (167.92) | 388.49 (201.62) | 351.76 (189.68) | | |
| Expository | Near | | | | | |
| | Premise | 369.68 (212.48) | 450.51 (244.79) | 409.16 (232.22) | 392.44 (213.67) | 391.28 (223.61) |
| | Inference | 330.33(180.54) | 423.28 (192.55) | 375.73 (192.08) | | |
| | Far | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| | Premise | 370.96 (246.68) | 451.80 (292.33) | 409.64 (272.30) | 390.09 (233.41) |
| | Inference | 343.40 (188.95) | 400.12 (175.59) | 370.54 (184.73) | |
| Total | | 335.55 (192.78) | 416.63 (223.54) | | |

**Table 8**

*Summary LMM for Premise and Inference Sentence Reading Times (Milliseconds per Word) Without Language Status as a Fixed Effect*

| Fixed effects | Estimated coefficient | *SE* | *t* | *p* |
|---|---|---|---|---|
| (Intercept) | 383.91 | 11.72 | 32.76 | <.001 |
| Reading strategy | -1.45 | 8.03 | -0.18 | .86 |
| **Word reading** | **-64.79** | **9.06** | **-7.15** | **<.001** |
| **Knowledge** | **-33.23** | **9.06** | **-3.67** | **<.001** |
| Text type | -18.09 | 9.16 | -1.98 | .06 |
| Distance condition | -2.62 | 2.18 | -1.20 | .23 |
| **Sentence type** | **-15.39** | **2.15** | **-7.14** | **<.001** |
| Reading strategy x Text type | 0.64 | 3.47 | 0.19 | .85 |
| Word reading x Text type | 6.24 | 3.83 | 1.63 | .11 |
| Knowledge x Text type | -2.83 | 3.85 | -0.73 | .46 |
| Reading strategy x Distance condition | -1.90 | 2.19 | -0.87 | .39 |
| Word reading x Distance condition | -0.16 | 2.43 | -0.07 | .95 |
| Knowledge x Distance condition | -1.62 | 2.48 | -0.65 | .51 |
| Text type x Distance condition | -2.26 | 2.18 | -1.04 | .30 |
| Reading strategy x Sentence type | -0.84 | 2.16 | -0.39 | .70 |
| Word reading x Sentence type | 2.45 | 2.41 | 1.02 | .31 |
| Knowledge x Sentence type | 0.23 | 2.45 | 0.09 | .93 |
| Text type x Sentence type | 3.12 | 2.15 | 1.45 | .15 |

| | | | | |
|---|---|---|---|---|
| Distance condition x Sentence type | 3.21 | 2.15 | 1.49 | .14 |
| Reading strategy x Text type x Distance | 2.79 | 2.19 | 1.27 | .20 |
| condition | | | | |
| Word reading x Text type x Distance | -1.48 | 2.43 | -0.61 | .54 |
| condition | 1.60 | 2.49 | 0.65 | .52 |
| Knowledge x Text type x Distance condition | | | | |
| Reading strategy x Text type x Sentence | 1.58 | 2.16 | 0.73 | .47 |
| type | 3.11 | 2.41 | 1.29 | .20 |
| Word reading x Text type x Sentence type | -1.09 | 2.45 | -0.44 | .66 |
| Knowledge x Text type x Sentence type | | | | |
| Reading strategy x Distance condition x | 0.99 | 2.16 | 0.46 | .65 |
| Sentence type | | | | |
| Word reading x Distance condition x | 0.36 | 2.41 | 0.15 | .88 |
| Sentence type | | | | |
| **Knowledge x Distance condition x** | **-5.05** | **2.45** | **-2.06** | **.04** |
| **Sentence type** | | | | |
| Text type x Distance condition x Sentence | 0.88 | 2.15 | 0.41 | .68 |
| type | 0.49 | 2.16 | 0.23 | .82 |
| Reading strategy x Text type x Distance | | | | |
| condition x Sentence type | 0.64 | 2.41 | 0.27 | .79 |
| Word reading x Text type x Distance | | | | |
| condition x Sentence type | 0.67 | 2.45 | 0.27 | .79 |

| | | | |
|---|---|---|---|
| Knowledge x Text type x Distance condition x Sentence type | | | |
| Random effects | | Variance | *SD* |
| Participant | (intercept) | 9115 | 95.47 |
| | Text type | 1006 | 31.71 |
| Item | (intercept) | 1733 | 41.63 |
| $R^2$ marginal[a] = 0.17, $R^2$ conditional[b] = 0.42 | | | |

*Note.* Observations = 5858; Participants = 152; Items = 24; $R^2$ calculated using the MuMIn package in R; Effects in bold are statistically significant. [a] Variance explained by the fixed effects. [b] Variance explained by the entire model including both fixed and random effects.

**Figure 1**

*Knowledge and Inference Question Accuracy*

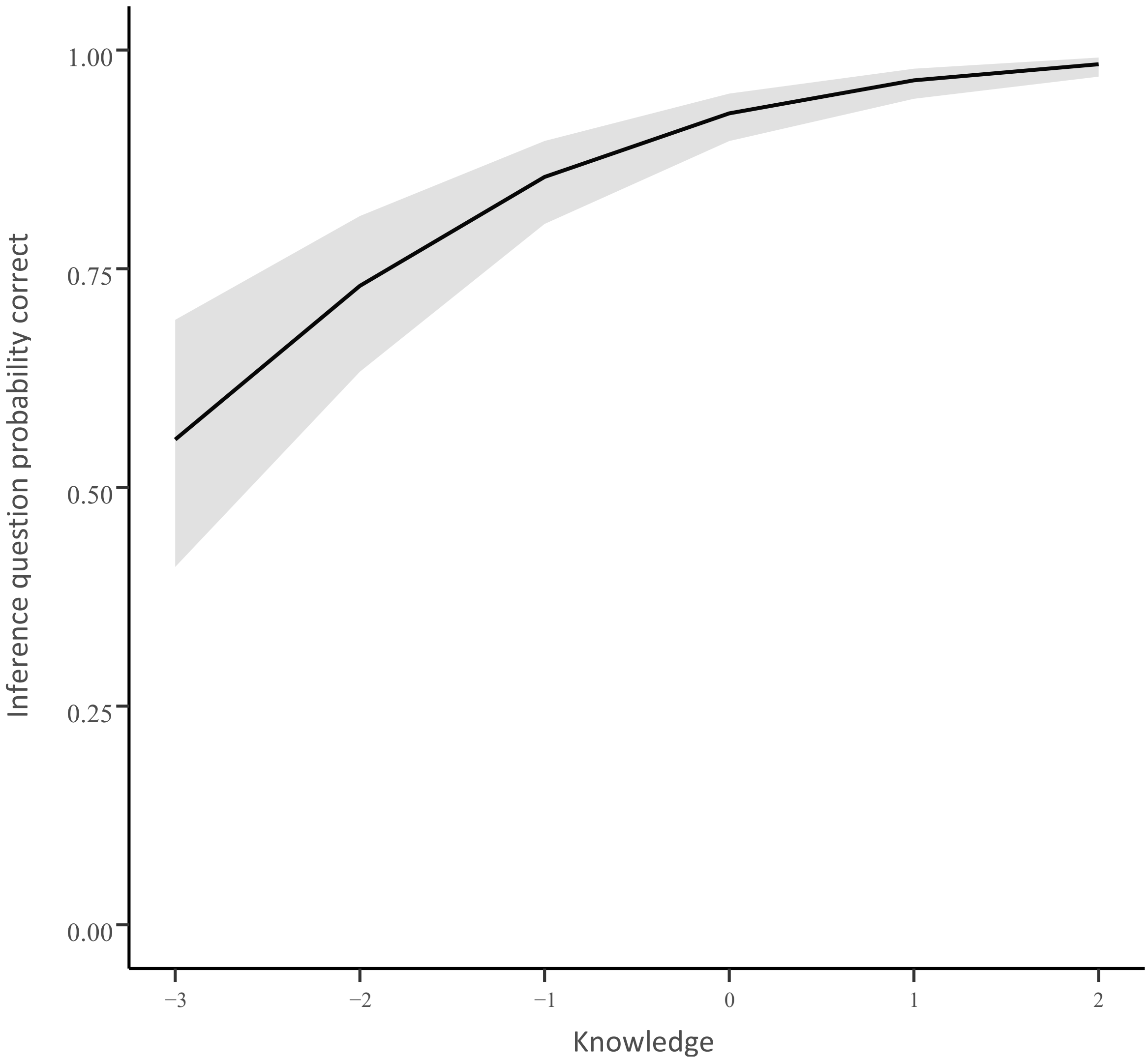

**Figure 2**

*Question Reading Time Interaction between Text Type and Word Reading*

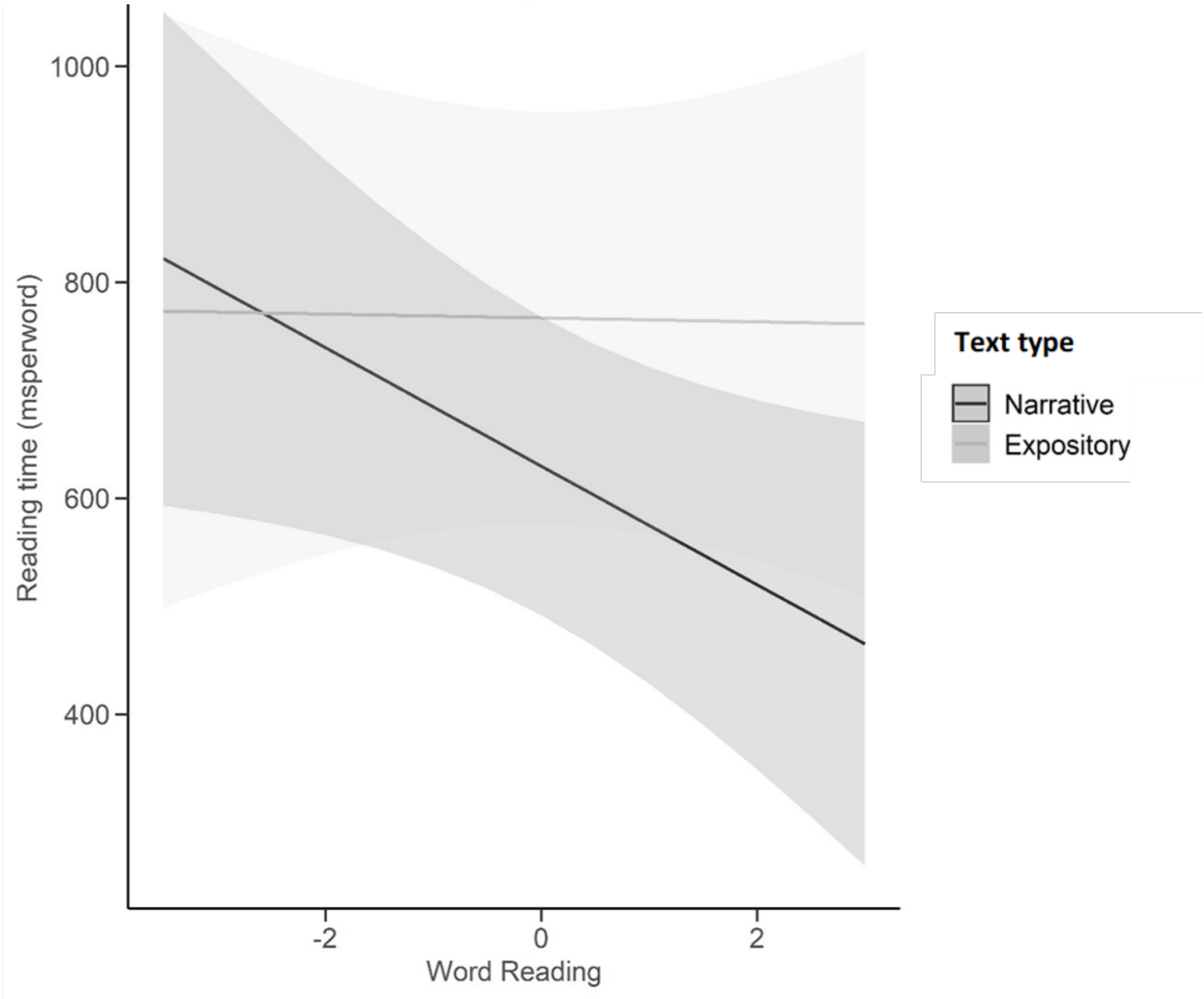

**Figure 3**

*Sentence Reading Time: Interaction between Knowledge, Distance Condition and Sentence Type*