Published in final edited form as:

Read Writ. 2015 May 1; 28(5): 587–609. doi:10.1007/s11145-014-9540-1.

Inferential Processing among Adequate and Struggling Adolescent Comprehenders and Relations to Reading Comprehension

Amy E. Barth,

University of Missouri-Columbia, Department of Special Education, 311B Townsend Hall, Columbia, MO 65211

Marcia Barnes,

University of Texas-Austin Department of Special Education, University of Texas-Austin, 1 University Station Stop D5300, Austin, TX 78712

David J. Francis,

University of Houston, Department of Psychology, Heyne Building 128C Houston, TX 77204

Sharon Vaughn, and

University of Texas- Austin, The Meadows Center for Preventing Educational Risk, 1912 Speedway SZB 228, Austin, TX 78712

Mary York

University of Houston, Texas Learning and Computational Center Annex, Room 225, Houston, TX 77204

Abstract

Separate mixed model analyses of variance (ANOVA) were conducted to examine the effect of textual distance on the accuracy and speed of text consistency judgments among adequate and struggling comprehenders across grades 6–12 ($n = 1203$). Multiple regressions examined whether accuracy in text consistency judgments uniquely accounted for variance in comprehension.

Results suggest that there is considerable growth across the middle and high school years, particularly for adequate comprehenders in those text integration processes that maintain local coherence. Accuracy in text consistency judgments accounted for significant unique variance for passage-level, but not sentence-level comprehension, particularly for adequate comprehenders.

Keywords

inferences; local and global coherence; reading comprehension; adolescent students

An element common to all the major theories of comprehension is the notion that the reader constructs a coherent mental representation or “mental model” of the situation described by text (Gernsbacher, 1990; Graesser, Singer, Trabasso, 1994; Johnson-Laird, 1983; Kintsch,

1998; van den Broek, Ridsen, Fletcher, & Thurlow, 1996). As the text proceeds in time, the situation model is revised with the integration of old and new meaning (Kintsch, 1988). Essential to this process of building a coherent situation model of text are inferential processes that support the integration of information within the text and between general knowledge and text (Graesser et al., 1994; McKoon & Ratcliff, 1992; van den Broek, Young, Tzeng, & Linderholm, 1999; van Dijk & Kintsch, 1983; Zwan & Radvansky, 1998).

Evidence from fluent adult readers suggests that semantic coherence, particularly local text coherence, is maintained through an unrestricted “search after meaning” process that involves mapping each proposition of an incoming sentence to propositions from the preceding sentence or two that are currently active in working memory (Graesser et al., 1994; O’Brien, Cook, & Geuraud, 2010; Rizzella & O’Brien, 2002). Global coherence may also be supported, in some cases, by these passive memory processes in which currently processed text “resonates” with or acts as a retrieval cue for text from an earlier processing cycle in long term working memory (O’Brien, Cook, & Peracchi, 2004; van den Broek, Rapp, & Kendeou, 2005); however, most often maintaining global coherence requires strategic mapping of incoming propositions to relevant general world knowledge and information encountered earlier in text that is beyond the span of working memory (Long & Chong, 2001). More active comprehension processes may also be engaged to repair coherence breaks at the local and global level and to meet a reader’s particular goals for reading, such as studying a text with the explicit goal of obtaining new knowledge (van den Broek et al., 2005).

The processes involved in maintaining local and global coherence during reading have been studied by manipulating features of the text, such as the distance in the text across which two sentences or ideas need to be integrated, concept repetition, degree of causal constraint provided by words, and so forth (review in Singer & Leon, 2007; Albrecht & O’Brien, 1991; Kintsch & Keenan, 1973; McKoon & Ratcliff, 1992; O’Brien et al., 2010). In terms of text distance, shorter distances between sentences-to-be-integrated may draw on processes that maintain local coherence, such as access to and retrieval of information from working memory (Albrecht & O’Brien, 1993; Singer, Andrusiak, Reisdorf, & Black, 1992). In contrast, larger textual distances are thought to tap processes that maintain global coherence, such as the integration of information and inferencing to bridge critical gaps (e.g., Long & Chong, 2001). In general, the greater the distance between two pieces of information in a text that must be integrated to maintain coherence, the less accurate and efficient readers are at executing these inferential processes (e.g. Albrecht & O’Brien, 1993; Long & Chong, 2001; Singer et al., 1992).

Text integration and inference processes have also been studied with respect to developmental and individual differences using methods that introduce various types of problems into texts (e.g., nonsense words, information violating general knowledge, or information inconsistent with prior text) and measuring children’s detection of these problems (Oakhill, Hartt, & Samols, 2005). Past research suggests that the ability to detect inconsistencies in text improves with age (Ackerman, 1984a; Baker, 1984; Pike, Barnes, & Barron, 2010; Zabrocky & Ratner, 1992); however, inconsistencies seem to be more difficult to detect over longer textual distances even in older elementary school-aged

children (Pike et al, 2010). The ability to detect inconsistencies in text is also related to comprehension skill (e.g., Long & Chong, 2001; Oakhill et al., 2005; Yuill, Oakhill, & Parkin, 1989). Skilled comprehenders are better overall at detecting inconsistent information in texts relative to less skilled comprehenders; however less skilled comprehenders, whether children or adults, seem to be particularly disadvantaged by increasing textual distance whether they are required to detect inconsistencies, make inferences, resolve lexical ambiguity, or use context to derive the meanings of new words (Barnes, Faulkner, Wilkinson, & Dennis, 2004; Cain, Oakhill, & Lemmon, 2004; Long & Chong, 2001; Oakhill et al., 2005; Yuill et al., 1989).

There are several unanswered questions about text integration processes that emerge from the extant literature reviewed above. Although there are some studies of text integration and inference in school-aged children and adults, developmental and individual difference studies of the ability to maintain local and global coherence in secondary school students are rare. Little is known about whether there are changes in text integration and inference processes in the secondary school years when students must increasingly use text as a source of new learning. As well, little is known about the pattern of individual differences in these abilities across the secondary school years; that is, by this age, are struggling comprehenders able to maintain local coherence, but have relatively more difficulty than adequate comprehenders in maintaining global coherence? Furthermore, except for one study using typically achieving 9th graders (Cromley & Azevedo, 2007) reporting that inference making had a small standardized direct effect (.19) on reading comprehension after controlling for various reading and reading-related skills, it is not known whether these text integration skills account for variability on standardized assessments of reading comprehension in the secondary school years (grades 6–12) over and above that accounted for by word reading accuracy, word reading fluency, working memory, and vocabulary, and whether such relations are moderated by comprehension skill.

To examine these questions, we employed a consistency detection task that measured 6th to 12th grade struggling and adequate comprehenders' ability to retrieve and integrate information in text by judging whether a new, continuation sentence was consistent or inconsistent with prior text. Text distance was manipulated to assess the effect of having to retrieve important elements of previously read text presented either immediately before (Near condition) or 5 sentences before (Far condition) the continuation sentence. In the far condition, the important prior text information is presumably in episodic memory and making a consistency judgment may serve to establish global coherence. In the near condition, prior text information may still be in working memory and the ability to make a consistency judgment may serve to establish local coherence. An inference was assumed to have been made only when the reader accepted a consistent continuation sentence or rejected an inconsistent continuation sentence.

Hypotheses

In terms of judging whether new information is consistent or inconsistent with the prior text, we hypothesized that continuation sentences would be judged more accurately and faster in the near versus the far condition regardless of grade or comprehension ability. Based on

prior research (Pike et al., 2010) among elementary grade children we expected that inconsistent continuations might be more difficult to reject in the far condition (be less accurate than consistent continuations) when information-to-be-integrated is separated in text. We also hypothesized that the inferential processing skills of struggling comprehenders would be less accurate and slower relative to adequate comprehenders particularly when information-to-be-integrated is separated in text, reflecting inefficient use of memory-based resources and/or lower standards of coherence that prevent the strategic deployment of fix-up strategies needed to maintain or repair global coherence.

We hypothesized that there may be grade-related changes in inferential processes across grades 6 through 12 (Ackerman & McGraw, 1991) for both adequate and struggling comprehenders, but particularly for those inferential processes that maintain global coherence; that is, in the far condition. Grade-related changes in inferential and text integration processes would provide support for continued developmental improvement across the secondary grades in these important text comprehension skills.

In terms of whether inferential processes predict reading comprehension, we hypothesized that inferential processes, because they involve between sentence integration, would account for unique variance for passage-level comprehension, but not single sentence-level comprehension after controlling for working memory and a number of other reading and cognitive variables. An additional research question was whether the relation of inferential processes to reading comprehension varies as a function of comprehension skill level; that is, is the relation of reading comprehension and the ability to integrate information across shorter and longer chunks of text moderated by comprehension skill level?

Method

Participants

School sites—This study was conducted in four school districts located in three cities and one suburb located in proximity to Houston, Texas. Students in grades 6–12 were recruited from two middle schools serving grades 6, three middle schools serving grades 7–8, and five high schools serving grades 9–12.

Criteria for participation—Participants were identified through a two-step process. First, we identified a large sample of struggling and adequate comprehenders ($n = 3,484$ struggling and $n = 10,953$ adequate comprehenders), based on their scores from the Texas Assessment of Knowledge Skills (TAKS; Texas Education Agency, 2004) administered in April of the previous school year. Students in the Struggling Comprehender group had obtained TAKS scaled scores ranging from 1,200–2,150. The Adequate Comprehender group obtained scale scores greater than 2,150. Students were excluded from participation if they were identified by their schools as Limited English Proficient (LEP), if their English Language Arts instruction was provided by a LEP teacher, or if they had a significant disability (e.g., intellectual-cognitive disabilities, severe behavioral disabilities, or autism).

Second, participants who consented and assented to participate in the study ($n = 929$ struggling and $n = 836$ adequate comprehenders), were administered the Woodcock

Johnson-III, Letter-Word Identification subtest (WJ-III: Woodcock, McGrew, & Mather, 2001). Students at or below the 20th percentile for their grade were excluded from further participation in order to obtain a sample with adequate word decoding.

The remaining students ($n = 597$ struggling and $n = 755$ adequate comprehenders) were then tested on a larger assessment battery that evaluated language, cognitive, and reading-related abilities. The sample was 51% male and 67% qualified for free or reduced lunch (with less than 1% not providing data on free-reduced lunch status). Demographic information is in Table 1.

Of the 1352 students who participated in the larger study, 136 students were excluded from this analysis because they were not administered the Bridging Inference Test due to absenteeism, computer failure, and scheduling conflicts. This resulted in a sample of 1216 students ($n = 541$ struggling comprehenders and $n = 675$ adequate comprehenders). An ANOVA confirmed that, on average, students who were administered the Bridge-IT did not differ significantly on the TAKS, $F(1,1350) = 3.00, p = .08$ or WJ-III Letter-Word Identification subtest, $F(1,1350) = .51, p = .48$ from those who were not administered the Bridge-IT or on demographic variables including gender, race/ethnicity, and reduced free lunch status: $\chi^2(1, N = 1352) = .76, p < .38$; $\chi^2(3, N = 1352) = 1.13, p = .77$; $\chi^2(1, N = 1352) = 1.33, p = .25$, respectively.

Testing procedures—Students were assessed by examiners who completed an extensive training program on test administration and scoring. Assessments were completed in quiet locations in the students' schools.

Measures

Measures to determine comprehension group membership

Texas Assessment of Knowledge and Skills (TAKS; Texas Educational Agency, 2010; 2011): The TAKS is a criterion-referenced assessment of reading comprehension that is specific for each grade that and aligned with grade-based reading standards and objectives. Students read expository and narrative texts and then answer multiple-choice, short-answer, and essay questions. Across grades 6–9 standard scores range from approximately 570.5 to 863.20; across grades 10–12 scores range from approximately 1953.70 to 2451.70. Because the scaling of the TAKS varies by grade (middle school versus high school) lexile scores are also reported which provide a common metric for comparison across grades 6–12. The internal consistency (coefficient alpha) of the 2010 TAKS for grades 7–12 range from .73 to .89; and 2011 TAKS alphas for grades 7–12 range from .87 to .89.

Woodcock-Johnson III Tests of Achievement (Woodcock et al. 2001): The Letter-Word Identification subtest is an individually administered assessment that measures children's ability to read real words. Reliabilities exceed .90 for secondary level students.

Measures of language, cognition, and reading

Kaufman Brief Intelligence Test- 2, Verbal Knowledge Subtest (K-BIT-2; Kaufman & Kaufman, 2004): This subtest is an individually administered test of receptive vocabulary

and general word knowledge. The participant is required to choose one of six illustrations that best corresponds to an examiner question. Internal consistency coefficients (split-half) for the verbal scores for grades 6–12 range from .89 to .94. Data on the Verbal Knowledge subtest are missing by design on half of the sample in order to reduce total assessment burden.

Woodcock-Johnson III Tests of Cognitive Abilities, Numbers Reversed Subtest

(Woodcock et al. 2001): This subtest assesses student's verbal working memory capabilities by asking students to repeat increasingly longer series of dictated digits in reversed order. Internal consistency exceeds .90 for secondary level students.

Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999):

The TOWRE is an individually administered assessment of reading fluency. For the Sight Word Efficiency (SWE) subtest participants read a list of 104 words as accurately and as quickly as possible; the number of words read correctly within 45 seconds is recorded. For the Phonemic Decoding Efficiency (PDE) subtest, participants read a list of 63 nonwords as accurately and as quickly as possible; the number of nonwords read accurately within 45 seconds is recorded. The standard score for the composite of SWE and PDE was used in analyses. For grades 6–12, the alternate form coefficients are .91 for SWE and .97 for PDE.

Test of Sentence Reading Efficiency and Comprehension (TOSREC; Wagner,

Torgesen, & Rashotte, 2010): The TOSREC is a 3-minute, group-based assessment of reading fluency and comprehension. Students silently read a series of short sentences and judge as quickly, but as accurately as possible, whether each is true or false. Average alternate-form coefficients for grades 6–12 range from .84 to .95 ($M = .90$).

Gates MacGinitie Reading Tests (MacGinitie, MacGinitie, Maria, Dryer, and Hughes,

2000): For this group-administered test of reading comprehension, students read short narrative and expository passages then answer questions assessing main idea, inference-making, and summarization. Form S was used (i.e., 6th grade form S; 7–9th grade form S; and 10–12th grade form S). Alternate-forms reliability coefficients range from .74–.89 across grades 6–12. The lexile score was used in the analyses.

Bridging Inferences Test (Bridge-IT):

The Bridge-IT was designed to measure the effect of textual distance (near versus far) (see Table 2 for example) on students' ability to judge whether a continuation sentence was consistent or inconsistent with prior text. The Bridge-IT comprises 32, five-sentence, narrative passages that were presented to students via Lenovo Thinkpad laptop computers. The E-Prime 2.0 Professional Network system was used to present the stimuli and collect all data.

Each five-sentence story was constructed to contain a key sentence of most importance for making the consistency judgment. In the near condition that key sentence (e.g., Alan does not like getting in trouble with his teacher) is the final sentence in the story; in the far condition it is the first sentence in the story. In both conditions, the information in the additional sentences of the story is compatible with either the consistent or inconsistent continuation sentence (as in Table 2). Making the correct judgment in the near condition

requires the reader to evaluate the information presented earlier in the story as well as the critical information presented in the final sentence of the story, which may still be in working memory. Making the correct judgment in the far condition requires the reader to evaluate the information they just read as well as the critical information presented in the first sentence of the story, which may need to be reactivated or retrieved from episodic memory.

The word “Ready” appeared on the computer screen followed one second later by a 5-sentence story. Students were instructed to press the spacebar when they had finished reading. Pressing the spacebar removed the story from the screen and an asterisk appeared in the middle of the screen to signal the presentation of the test sentence. Participants were told to read the test sentence and then to press the green button if they thought the sentence was a good continuation of the story (“consistent”) and the red button if they thought it was not a good continuation (“inconsistent”). They were instructed to make their decisions as quickly but as accurately as possible. Two practice items were used to explain to students the difference between good and poor continuations in relation to judging consistency between the story and the test sentence, and to familiarize them with the procedure.

Students received a testlet that comprised 8 near-consistent, 8 far-consistent, 8 near-inconsistent, and 8 far-inconsistent items, with items counterbalanced across conditions. For each item, reading time was obtained for the passage and accuracy and response time (RT) were obtained for continuation sentence judgments. Continuation sentences were 3 to 12 words in length across items. Word length was equivalent for consistent and inconsistent versions of the continuation sentence for a particular passage.

A total accuracy score and 4 condition accuracy scores (i.e., near-consistent, far-consistent, near-inconsistent, and far-inconsistent) were calculated. All accuracy scores represented the proportion of items answered correctly after trimming for outliers as follows: Passage reading times greater than 15 words per second (WPS) and continuation RTs greater than 10 WPS were considered pre-pushes suggesting that the student did not read the material; passage reading times less than 1 WPS and continuation RT less than .5 WPS were considered extremely slow reading for secondary level students who had been screened to eliminate decoding problems, suggesting that either a technical problem was encountered or the score was an outlier. A minimal amount of data were considered pre-pushes or outliers and trimmed from the data set: near-consistent ($n = 7$); near-inconsistent ($n = 19$); far-consistent ($n = 18$); far-inconsistent ($n = 32$). Trimmed data were excluded from further analyses; however, results are the same with trimmed data included in the analyses. Six percent of students were missing one or more of the condition scores after applying these rules.

Mean continuation RT was computed for each individual for each of the four conditions regardless of the accuracy of student’s response. Incorrect items were included because accuracy of the student’s response to the prompt was not expected or guaranteed given the task complexity.

Average reliability coefficients (Kuder-Richardson 20) are .85 for near-consistent; .87 for near-inconsistent; .83 for far-consistent; and .87 for far-inconsistent continuations.

Analytic Plan

To determine the effect of textual distance (near or far) and text consistency (consistent or inconsistent continuation) on accuracy and rate of sentence judgments and whether these effects depended on students' grade level (grades 6–12) or reader subgroup (struggling or adequate comprehender), separate four-way, textual distance x text consistency x reader subgroup x grade ($2 \times 7 \times 2 \times 2$) mixed model analyses of variance (ANOVA) were conducted for accuracy and rate of sentence judgments, with reader sub-group and grade as between subjects factors and textual distance and text consistency as within-subjects factors. We allowed for random effects of students, but ignored the clustering of students within schools for several, related reasons. First, the number of schools was relatively small for estimating random effects of schools. Moreover, the variance attributable to schools was very small. Specifically, unconditional, interclass correlations for the outcome measures ranged from .005 to .023 for the speed and accuracy measures, suggesting that the degree of non-independence in the data is very small. The small magnitude of the ICCs is not surprising given that the outcome measure is a laboratory measure and not directly tied to instruction in the schools.

For each mixed model ANOVA (accuracy and rate), the following covariance structures were compared to determine which structure most accurately characterized the structure of the covariance matrix. This step was undertaken to safeguard against the effects of using an incorrectly specified covariance structure on estimation of the standard errors of the experimental conditions, which would typically result in an inflated Type I error rate. The different covariance structures estimated were (a) unstructured, (b) compound symmetry, (c) compound symmetry taking into account subgroup and (d) compound symmetry taking into account grade. To determine which covariance structure resulted in the best fitting model three indices were examined: Akaike's Information Criteria (AIC), Akaike's Information Criteria Corrected (AICC), and Bayesian Information Criteria (BIC). Each of the indices is a function of the log likelihood and can be compared across models with smaller values of the indices suggesting better fitting models.

The Tukey test, a multiple comparison procedure, was used after obtaining a significant omnibus test result to correct for family-wise error rate. Effect sizes are reported using eta-squared, which describes the ratio of variance explained in the dependent variable by an independent variable while controlling for other predictors. Eta-squared represents the ratio of the between groups sum of squares to the total sum of squares and is analogous to r^2 . As a general rule of thumb, an effect size of 1% is considered small, 10% medium, and 25% large (Cohen, 1968).

The second set of analyses used multiple regression. In regression models 1 and 2 (predicting passage-level and sentence-level comprehension, respectively), we entered gender, reduced free lunch status, race, grade level, word reading accuracy, word reading efficiency, working memory, vocabulary knowledge, group status, accuracy in the Near condition, and finally the product of reader group status and near condition accuracy.

Regression models 3 and 4 (predicting passage-level and sentence-level comprehension, respectively) were identical to models 1 and 2 except that accuracy in the Far condition was entered instead of near condition accuracy. Separate regression models using Bridge-It Near versus Far accuracy were conducted because of the pattern of findings for the two text distances reported below.

Results

Preliminary Analyses

Of the 1,216 students who were administered the Bridging Inferences task, 13 students were dropped due to reliability issues at the item level, which resulted in a final sample of 1,203 students.

We evaluated distributional data both statistically and graphically for skewness, kurtosis, and normality. The following scores possessed high skew: near-consistent accuracy (1.20) and near-inconsistent accuracy (−1.36). The following scores possessed high kurtosis: near-consistent RT (21.28) and far-consistent RT (3.90). The following academic variables were missing some data: TOWRE Composite ($n = 2$); WJ-III Numbers Reversed ($n = 4$); TOSREC ($n = 27$), and Gates MacGinitie Passage Comprehension ($n = 18$). As mentioned above, the KBIT-2 Verbal Knowledge was missing by design on approximately half the participants ($n = 563$). Eight percent of students were missing one or more academic or Bridge-IT test scores used in the analyses, not including the KBIT-2 Verbal Knowledge.

Table 3 shows demographic information for each reader subgroup. Chi-square analyses and ANOVAs were conducted to determine if reader subgroups differed on demographic indices. Struggling and adequate comprehenders did not differ by age, $F(1, 1201) = .96$, $p = .33$; but differed significantly by gender, race/ethnicity, and reduced free lunch status: $\chi^2(1, N = 1203) = 25.35$, $p < .0001$; $\chi^2(3, N = 1203) = 12.27$; $p = .007$; $\chi^2(1, N = 1203) = 28.68$, $p < .0001$, respectively. Struggling comprehenders had a significantly higher proportion of males (60% versus 45%), Hispanics (56% versus 47%), and students receiving reduced free lunch (74% versus 59%) relative to adequate comprehenders. Because of these differences, free reduced lunch, gender, and race were included as covariates in the analyses below to obtain the best estimate of the magnitude of the difference that is due to reader subgroup status and not the demographic variables that are correlated with reader subgroup status.

Table 4 reports the means and standard deviations for the study measures by reader subgroup.

Accuracy

A $2 \times 7 \times 2 \times 2$ (Reader Subgroup x Grade Level x Textual Distance x Text Consistency) mixed model ANOVA, with Textual Distance and Text Consistency as within-subjects actors, was conducted. In the full model, with all the possible 2-way, 3-way, and 4-way interactions, the following main effects were significant: reader subgroup, $F(1, 1183) = 155.35$, $p < .0001$, $\eta^2 = .116$; grade, $F(6, 1183) = 155.25$, $p < .0001$, $\eta^2 = .441$; textual distance, $F(1, 1185) = 952.22$, $p < .0001$, $\eta^2 = .446$; and text consistency, $F(1, 1187) = 48.86$,

$p < .0001$, $\eta^2 = .04$. The main effects were qualified by a significant three-way interaction among reader subgroup, textual distance, and text consistency, $F(1, 1168) = 4.38$, $p = .04$, $\eta^2 = .004$; and a significant two-way interaction between grade and textual distance, $F(6, 1185) = 3.65$, $p = .0013$, $\eta^2 = 0.018$.

The significant three-way interaction was followed by tests of simple interaction effects, which, in turn, were followed up by simple-simple effects, where necessary. Tests of simple interaction effects showed that there was a significant interaction between reader subgroup and text consistency in the near condition $F(1, 1168) = 3.88$, $p = .0498$, $\eta^2 = .003$, but not in the far condition, $F(1, 1168) = .96$, $p = .3266$, $\eta^2 = .0008$. Further, examination showed that in the near condition, adequate comprehenders were more accurate when the information-to-be integrated was inconsistent with the preceding sentence than when it was consistent $t(668) = -2.13$, $p = .03$ (see Table 5). In contrast, struggling comprehenders performed similarly regardless of consistency $t(530) = 1.25$, $p = .21$. In the far condition, continuation sentences consistent with the prior text were judged more accurately than inconsistent continuation sentences by both adequate comprehenders $t(661) = 8.36$, $p < .001$ and struggling comprehenders $t(521) = 5.55$, $p < .001$.

The interaction of textual distance and grade was followed up by testing the simple effects of grade within distance. The simple effect of grade within distance was examined to determine whether the effect of condition (near vs. far) differs across grades. In the near condition, the effect of grade within distance was significant $F(6, 1194) = 13.25$, $p < .0001$, $\eta^2 = .06$. Students in grades 6 and 7 were less accurate than students in grades 10–12; and students in grades 8 and 9 were less accurate than students in grade 10 (all $ps < .007$) (see Table 6). In the far condition, the simple effect of grade within distance was also significant $F(6, 1189) = 4.47$, $p = .0002$, $\eta^2 = .02$. Students in grade 10 were more accurate than students in grades 6–9 ($ps < .007$), and students in grade 12 were more accurate than students in grade 9 ($p < .007$). Results suggest fewer grade-related changes for far inference than for near inference as can be seen in Table 6 and Figure 1.

Response Time

A $2 \times 7 \times 2 \times 2$ (Reader Subgroup x Grade x Textual Distance x Text Consistency) mixed model ANOVA, with Textual Distance and Text Consistency as within-subjects factors, was conducted. In the full model, with all the possible 2-way, 3-way, and 4-way interactions, the following main effects were significant: reader subgroup, $F(1, 1182) = 86.47$, $p < .0001$, $\eta^2 = .068$; grade level, $F(6, 1182) = 30.12$, $p < .0001$, $\eta^2 = .133$; text consistency $F(1, 1181) = 127.91$, $p < .0001$, $\eta^2 = .098$; and textual distance $F(1, 1181) = 35.63$, $p < .0001$, $\eta^2 = .029$, such that sentence continuation judgments were faster in the near than in the far condition (2.39 vs. 2.33 WPS). The main effects of reader subgroup and text consistency were qualified by a significant interaction between reader subgroup and text consistency, $F(1, 1181) = 11.91$, $p = .0006$, $\eta^2 = .010$.

Tukey tests on the main effect of grade showed that students in grade 6 were slower at sentence continuation judgments than students in grades 8–12; students in grade 7 were slower than in grades 9–12; students in grade 8 were slower than in grades 10–12; and students in grade 9 were slower than in grades 11 and 12 (all $ps < .007$). The means and

standard deviations (*SD*) of sentence continuation judgments measured in words per second from grades 6–12 were: 1.94 (.55); 2.07 (.52); 2.22 (.62); 2.36 (.64); 2.55 (.68); 2.57 (.71); and 2.66 (.70), respectively.

The simple effect of text consistency within reader group was examined to understand the interaction. Consistent continuations were judged more slowly than inconsistent continuations by both adequate comprehenders and struggling comprehenders, ($t(669) = -8.48, p < .0001$ and $t(530) = -4.66, p < .0001$, respectively; however, the difference between consistent and inconsistent continuations was relatively greater among adequate comprehenders than struggling comprehenders. For adequate comprehenders, the mean (standard deviation) for words read per second were 2.43 (.73) and 2.63 (.75) for consistent and inconsistent continuations, respectively. For struggling comprehenders, the means (standard deviation) were 2.09 (.65) and 2.19 (.64) for consistent and inconsistent continuations, respectively. These effects were small.

Contribution of Bridging Inference to Reading Comprehension

In the first model, with TOSREC – Standard Score regressed on grade level, WJ-III Letter Word Identification, TOWRE Composite, WJ-III Numbers Reversed, KBIT-2 Verbal Knowledge, Bridging Inferences Near accuracy, group status, and the product of group status and Bridging Inferences Near accuracy, 48% of the variance was accounted for $F(8, 603) = 69.23, p < .0001$. The TOWRE, WJ-III Numbers Reversed, KBIT-2 Verbal Knowledge, and Bridge-It-Near each uniquely accounted for 9%, 1%, 12%, and .7% of the variance in sentence-level comprehension, respectively, after controlling for grade level (Table 7).

In the second model, with Gates MacGinitie Reading Tests – Lexile Score regressed on the same variables as the model above, 55% of the variance was accounted for $F(6, 614) = 93.33, p < .0001$. WJ-III Letter Word Identification (17%), WJ-III Numbers Reversed (1%), KBIT-2 Verbal Knowledge (8%), Bridge-IT-Near (3%), and the interaction of Bridge-It-Near and group status (.5%) each accounted for unique variance in passage-level comprehension after controlling for grade level (Table 8). Separate regression models indicated that Bridging Inferences-Near uniquely accounted for 5% of the variance in reading comprehension among adequate comprehenders $F(6, 345) = 54.47, p < .0001, R^2 = .49$ and 2% among struggling comprehenders $F(6, 264) = 42.16, p < .0001, R^2 = .49$.

In the third model, with TOSREC – Standard Score regressed on grade level, WJ-III Letter Word Identification, TOWRE Composite, WJ-III Numbers Reversed, KBIT-2 Verbal Knowledge, Bridging Inferences Far accuracy, group status, and the product of group status and Bridging Inferences Far accuracy, 48% of the variance was accounted for $F(8, 603) = 68.21, p < .0001$. The TOWRE uniquely accounted for 9% and KBIT-2 Verbal Knowledge uniquely accounted for 12% of the variance in sentence-level comprehension after controlling for grade level (Table 9).

In the fourth model, with Gates MacGinitie Reading Tests – Lexile Score regressed on the same variables as the model above, 54% of the variance was accounted for $F(6, 614) = 89.95, p < .0001$. WJ-III Letter Word Identification (17%), WJ-III Numbers Reversed (1%),

KBIT-2 Verbal Knowledge (8%), and the interaction of reader sub-group status and Bridge-IT-Far (.5%) uniquely accounted for variance in passage-level comprehension after controlling for grade level (Table 10). Separate regression models indicated that Bridging Inferences-Far uniquely accounted for 4% of the variance in reading comprehension among adequate comprehenders $F(6, 345) = 53.50, p < .0001, R^2 = .48$ but did not account for unique variance among struggling comprehenders $F(6, 264) = 39.47, p < .0001, R^2 = .47$. Among adequate comprehenders, with Gates McGinitie Reading Test –Lexile Score regressed on grade, WJ-III Letter Word Identification, TOWRE, WJ-III Numbers Reversed, KBIT Verbal Knowledge, Bridge-IT Near and Bridge-IT Far, Bridge-It Far uniquely accounted for 3% of the variance after controlling for all variables in the model, $F(7, 343) = 51.10, p < .0001, R^2 = .51$.

Discussion

The main findings of this study are that: (a) adequate comprehenders and older adolescents are more accurate and faster at judging the consistency of a test sentence when information-to-be-integrated was adjacent in text versus separated in text by several sentences; (b) the largest grade-related changes in making these types of judgments were in the near condition suggesting considerable growth across the secondary school years in text integration skills that affect the maintenance of local coherence; and (c) the ability to integrate information across short distances accounts for unique variance in both sentence-level and passage-level comprehension whereas the ability to integrate information across longer distances accounts for unique variance in passage-level but not sentence-level comprehension, and only among adequate comprehenders.

Grade-related Differences in Inferential Processing

Results of this study show that the greatest grade-related changes in inference-making accuracy and speed were observed when the information-to-be integrated was between adjacent sentences. When information-to-be-integrated was separated in text, fewer grade related differences in the rate and accuracy of consistency detection were observed. These findings extend prior research on age-related differences in inferential processing among elementary school students (Cain, Oakhill, & Elbro, 2003; Casteel & Simpson, 1991; Pike et al., 2010) and show that there is considerable continued growth in text integration skills well into the high school years. However, these grade-related gains appear to be greatest for maintaining local rather than global coherence. To our knowledge, this is the only study looking at grade-related differences in text integration skills at the local and global levels across the secondary school years. The findings suggest that processes related to constructing local coherence are not yet fully established in middle school students, which may have implications for comprehension instruction for this age group.

Individual Differences in Inferential Processing

Both adequate and struggling comprehenders were more accurate and faster in their ability to integrate information in the near versus far inference condition; however, adequate comprehenders were more accurate and faster at judging the consistency of continuation sentences with prior text relative to struggling comprehenders in both conditions. Based on

prior studies (e.g., Cain, Oakhill, & Bryant, 2004; Oakhill, Hart, & Samols, 2005), and given the age of our sample, we expected that the largest differences between adequate and struggling comprehenders might emerge for judgments in the far versus the near condition. However, similar to the grade-related findings, large differences between adequate and struggling comprehenders were also found on items involving judgments of text consistency between adjacent sentences.

In the far integration condition the performance of the struggling comprehenders was close to chance (53% and 56% accuracy in 6th and 12th grades, respectively), suggesting the presence of floor effects particularly for far-inconsistent items. Among adequate comprehenders, performance was also low but above chance levels (58% to 66%, between 6th and 12 grades) with higher overall performance on far-consistent than far-inconsistent items. Both groups had relatively greater difficulty rejecting incorrect continuations across larger versus shorter textual distances. Although this is consistent with the idea that when information required for making a judgment is no longer in working memory decisions may be made on the effects for the far inference items in this study limit how well we can estimate the global coherence processes in secondary school struggling comprehenders.

What might account for low performance on the far inference items? According to memory-based accounts, coherence breaks may be caused by difficulties in either the activation or integration of information. In the activation phase, passive search after meaning processes seek to reactivate outdated information, with reactivation influenced by the repetition of concepts and one's knowledge of the topic (Magliano & Radvansky, 2001; O'Brien et al., 1995; O'Brien et al., 2010). However, even when prior text and related information is sufficiently activated, violations of global coherence may still not be detected (Long & Chong, 2001; Singer et al., 2001). In the current paradigm, it is worth noting that key information from the text (i.e., that Alan does not like to get into trouble with his teacher) is mentioned only once, which may reduce the probability that this information is activated in working memory and therefore available to the reader to support integration. Thus, these constructed texts may differ in important ways from typical text, partly accounting for low accuracy in the far condition. Differences in the cohesiveness of the texts in the near and far conditions could have been induced by the change in placement of the key sentence in the passage, and this change in text cohesion might have affected readers' accuracy. However, analysis of the passages using cohmetrix (Graesser, McNamara, & Kulikowich, 2011) did not show significant differences on several indices of cohesiveness between the near and far conditions. Nevertheless, despite a lack of statistical differences in cohesiveness between passages in near and far conditions, we cannot rule out the possibility that differences in text cohesion contribute to the distance effect observed in the present study. The possible effects of cohesiveness are likely best investigated by modeling the simultaneous effects of both person- and text-characteristics on inference performance. Clearly, text distance significantly affected readers' ability to reconcile information between competing models of the situation described in text (e.g., that Alan does not like to get into trouble with his teacher, but also that the other boys told a funny joke and are laughing) for the types of texts used in the current study (Gernsbacher & Faust, 1991; Pike et al., 2010) even if the mechanism by which distance affected readers' comprehension cannot be unambiguously determined from the current experiment and analyses.

In sum, the adolescent struggling comprehenders in this study demonstrated difficulties relative to adequate comprehenders particularly in the ability to maintain both local coherence. The difficulties of these struggling adolescent comprehenders in maintaining local coherence will affect the accuracy and quality of their text-based and situational representations.

Contribution of Inferential Processing to Reading Comprehension

Performance on the Bridge-IT (Near and Far) uniquely predicted sentence-level and passage-level comprehension after controlling for decoding, language, and cognitive variables. These findings converge with past research suggesting that inferential processing plays an important role in reading comprehension among elementary grade readers (Cain, Oakhill, & Bryant, 2004; Pike et al., 2010) and also secondary grade readers (Cromley & Azevedo, 2007; Karasinski & Ellis Weismer, 2010). The unique contributions of the present study are: (a) inferential processing at the local level was related to single sentence comprehension (although the unique contribution of inference to sentence-level comprehension was small); (b) inferential processing at the local and global level uniquely contributed to passage-level comprehension; and (c) the relation of inferential processing to reading comprehension was moderated by comprehension skill.

Among struggling comprehenders, accuracy on near inferences uniquely accounted for variance in passage-level comprehension for both adequate comprehenders (5% and 2%, respectively). In contrast, inferential processing at the global level (accuracy on far inferences) uniquely accounted for 4% of the variance in reading comprehension but only for adequate comprehenders. Additionally, among adequate comprehenders, inferential processing at the global level uniquely accounted for 3% of the variance in passage-level comprehension after controlling for inferential processing at the local level. Adequate comprehenders may be more likely to engage both strategic processes to maintain or repair coherence and more passive processes to link up information in text and information in text with background knowledge (van den Broek et al., 2005) relative to struggling readers, particularly when critical information to-be-integrated is separated in text. The findings could also suggest that factors other than text integration are more important for accounting for comprehension in struggling adolescent comprehenders such as the accessibility of relevant information from the text (Vanderwood, McGrew, Flanagan, & Keith, 2002), or lack of critical word and world knowledge (Compton, Miller, Elleman, & Steacy, 2014), which has both direct and indirect effects (through inference-making) on reading comprehension (Cromley & Azevedo, 2007).

Limitations and Future Directions

A limitation of the study is that the items in the far condition, particularly far-inconsistent, exhibited floor effects for both struggling comprehenders and perhaps for some adequate comprehenders as well. Rarely do narrative texts purposely require the reader to resolve two competing mental models with five sentences of text where there is little development of the characters and the passages do not include causal-explanation sentences to refute outdated or inconsistent information (Kendeou, Smith, & O'Brien, 2013). Future research should further examine issues related to the maintenance of global cohesion but do so with narrative texts

that more closely align with the type of materials read by school age children in the secondary grades.

A second limitation of the study is that we are unable to determine whether the difficulties in inferential processing at the local versus global level are caused by problems in the activation of information or the integration of information. Although research conducted with adult readers (e.g., Singer et al., 2001) indicates that difficulties in establishing global coherence are largely caused by problems integrating information, it is beyond the scope of this study to determine if a similar pattern exists among adequate and struggling secondary level comprehenders. We were also unable to determine what types of relational information in text struggling comprehenders have most difficulty integrating to maintain local coherence. For example, the study did not compare the ability to maintain local (or global) coherence with respect to pronominal reference, or for information involving physical causal relations or character goals and motivations, and so forth.

Results of the current study indicate that older struggling comprehenders have difficulty making inferences even when the relevant information has just been read. Knowing about these difficulties in maintaining local coherence is relevant for thinking about the level at which interventions need to be aimed for these older struggling readers. Instruction in inference-making may need to focus on how to integrate relevant information within and between sentences and across short chunks of text to facilitate the construction of a locally coherent text-based representation. Intervention studies have typically not manipulated inference-making between back-to-back sentences as a springboard for inference-making across larger chunks of text nor have they manipulated the number of practice trials required to attain mastery at various passage lengths. Future research of this sort may help determine how to best instruct students to consistently integrate information and make inferences to establish a more locally and globally coherent situation model of the text (Kintsch, 1998).

Summary

The findings from the current study provide new information about inference and text integration processes through the secondary school years and suggest considerable change in the efficiency of those integration processes that contribute to establishing local coherence. The findings also provide information on the marked difficulties of struggling adolescent comprehenders in integrating information across larger textual distances and the less anticipated finding that these students have difficulties, compared to their adequately comprehending peers, in establishing local semantic coherence. To the extent that the situation models of text for struggling adolescent comprehenders lack both local and global coherence, they will be unable to support adequate levels of comprehension and new learning. Finally, the current findings show that inferential processes are related to sentence-level and passage-level comprehension, and that this relationship is moderated by comprehension skill at the passage-level.

Acknowledgments

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305F100013 to the University of Texas-Austin as part of the Reading for Understanding Research Initiative as well as grant K08 HD068545-01A1 from the Eunice Kennedy Shriver National Institute of Child Health and

Human Development (NICHD). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The authors wish to acknowledge the invaluable contributions of the middle and high school staff and students in Channelview ISD, Dickinson ISD, Galveston ISD and Humble ISD in Texas and the assistance of Maria Hernandez, Sharon Kalinowski, Frances Leal, and Catherine Watkins.

References

- Ackerman BP. The effects of storage and processing complexity on comprehension repair in children and adults. *Journal of Experimental Child Psychology*. 1984a; 37:303–334.
- Ackerman BP, McGraw M. Constraints on the causal inferences of children and adults in comprehending stories. *Journal of Experimental Child Psychology*. 1991; 51:364–394.
- Albrecht JE, O'Brien EJ. Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1993; 19:1061–1070.
- Albrecht JE, O'Brien EJ. Effects of centrality on retrieval of text-based concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1991; 17:932–939.
- Baker L. Spontaneous versus instructed use of multiple standards for evaluating comprehension: Effects of age, reading proficiency, and type of standard. *Journal of Experimental Child Psychology*. 1984; 38:289–311.
- Barnes MA, Faulkner H, Wilkinson M, Dennis M. Meaning construction and integration in children with hydrocephalus. *Brain and Language*. 2004; 89:47–56. [PubMed: 15010236]
- Cain K, Oakhill J, Bryant P. Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*. 2004; 96:31–42.
- Cain K, Oakhill J, Elbro C. The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*. 2003; 30:681–694. [PubMed: 14513473]
- Cain K, Oakhill J, Bryant P. Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*. 2004; 96:31–42.
- Casteel MA, Simpson GB. Textual coherence and the development of inferential generation skills. *Journal of Research in Reading*. 1991; 14:116–129.
- Cohen J. Multiple regression as a general data-analytic system. *Psychological Bulletin*. 1968; 70:426–443.
- Compton D, Miller A, Elleman A, Steacy L. Have we forsaken reading theory in the name of “quick fix” interventions for children with reading disability? *Scientific Studies of Reading*. 2014; 18:55–73.
- Cromley JG, Azevedo R. Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*. 2007; 99:311–325.
- Gernsbacher, MA. *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1990.
- Gernsbacher MA, Faust ME. The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1991; 17:245–262.
- Graesser AC, McNamara D, Kulikowich J. Coh-metrix: Providing multi-level analyses of text characteristics. *Educational Researcher*. 2011; 40:223–234.
- Graesser AC, Singer M, Trabasso T. Constructing inferences during narrative text comprehension. *Psychological Review*. 1994; 101:371–395. [PubMed: 7938337]
- Johnson-Laird, PN. *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, UK: Cambridge University Press; 1983.
- Karasinski C, Ellis-Weismer S. Comprehension of inferences in discourse processing by adolescents with and without language impairments. *Journal of Speech, Language, and Hearing Research*. 2010; 53:1268–1279.
- Kaufman, AS.; Kaufman, NL. *Kaufman Brief Intelligence Test, Second Edition (K-BIT-2)*. Minneapolis, MN: Pearson Assessment; 2004.

- Kendeou P, Smith E, O'Brien E. Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2013; 39:854–865.
- Kintsch W. The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review*. 1988; 95:163–182. [PubMed: 3375398]
- Kintsch, W. *Comprehension: A paradigm for cognition*. New York: Cambridge University Press; 1998.
- Kintsch W, Keenan J. Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*. 1973; 93:411–428.
- Long DL, Chong JL. Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2001; 27:1424–1429.
- MacGinitie, WH.; MacGinitie, RK.; Dreyer, LG. *Gates-MacGinitie Reading Test*. 4. Itasca, IL: Riverside Publishing Company; 2000.
- Magliano J, Radvansky G. Goal coordination in narrative comprehension. *Psychonomic Bulletin and Review*. 2001; 8:372–376. [PubMed: 11495128]
- McKoon G, Ratcliff R. Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology*. 1992; 18:1155–1172. [PubMed: 1447546]
- Oakhill J, Hartt J, Samols D. Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing*. 2005; 18:657–686.
- O'Brien EJ, Albrecht JE, Hakala CM, Rizzella ML. Activation and suppression of antecedents during reinstatement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1995; 21:626–634.
- O'Brien EJ, Cook AE, Gueraud S. Accessibility of outdated information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2010; 36:979–991.
- O'Brien E, Cook AE, Peracchi KA. Updating situation models: Reply to Zwan and Madden (2004). *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2004; 18:777–784.
- Pike MM, Barnes MA, Barron RW. The role of illustrations in children's inferential comprehension. *Journal of Experimental child Psychology*. 2010; 105:342–255.
- Rizzella ML, O'Brien EJ. Retrieval of concepts in script-based texts and narratives: The influence of general world knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2002; 28:780–790.
- Singer M, Andrusiak P, Reisdorf PE, Black NL. Individual differences in bridging inferences processes. *Memory and Cognition*. 1992; 20:539–548. [PubMed: 1453971]
- Singer, M.; Leon, J. Psychological studies of higher language processes: Behavioral and empirical studies. In: Schmalhofer, F.; Perfetti, C., editors. *Higher level language processes in the brain: Inference and comprehension processes*. Mahwah, NJ: Lawrence Erlbaum Associates; 2007. p. 9-25.
- Singer M, Kintsch W. Text retrieval: A theoretical exploration. *Discourse Processes*. 2001; 31:27–59.
- Texas Educational Agency. [on December 19 2007] TAKS: Texas Assessment of Knowledge and Skills. Information booklet: Reading. 2004b. grade 6 – Revised Author. Retrieved from <http://www.tea.state.tx.us/student.assessment/taks/booklets/reading/g6e.pdf>
- Torgesen, J.; Wagner, R.; Rashotte, C. *Test of Word Reading Efficiency*. Austin, TX: Pro-Ed; 1999.
- Van den Broek P, Rapp D, Kendeou P. Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes*. 2005; 39:299–316.
- van den Broek, P.; Ridsen, K.; Fletcher, C.; Thurlow, R. A 'landscape' view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In: Bruce, B.; Graesser, A., editors. *Models of understanding text*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1996. p. 165-187.
- van den Broek, P.; Young, M.; Tzeng, Y.; Linderholm, T. The landscape model of reading: Inferences and the online construction of a memory representation. In: Goldman, SR.; van Oostendorp, H., editors. *The construction of mental representations during reading*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 1999. p. 71-98.

- Vanderwood ML, McGrew KS, Flanagan DP, Keith TA. The contribution of general and specific cognitive abilities to reading achievement. *Learning and Individual Differences*. 2002; 13:159–188.
- van Dijk, TA.; Kintsch, W. *Strategies and discourse comprehension*. New York, NY: Academic Press; 1983.
- Wagner, RK.; Torgesen, J.; Rashotte, C.; Pearson, N. *Test of Sentence Reading Efficiency (TOSRE)*. Austin, TX: PRO-ED, Inc; 2010.
- Woodcock, RW.; McGrew, KS.; Mather, N. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing; 2001.
- Yuill N, Oakhill JV, Parkin AJ. Working memory, comprehension ability, and the resolution of text anomaly. *British Journal of Psychology*. 1989; 80:351–361. [PubMed: 2790393]
- Zabracky K, Ratner H. Effects of passage type on comprehension monitoring and recall in good and poor readers. *Journal of Reading Behavior*. 1992; 24:373–391.
- Zwaan RA, Radvansky GA. Situation models in language comprehension and memory. *Psychological Bulletin*. 1998; 123:162–185. [PubMed: 9522683]

Figure 1.
Accuracy Data - Simple Effect of Grade within Distance

Table 1**Demographics by Grade and Reader Group**

	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Struggling Comprehender Group							
<i>n</i>	77	80	104	83	100	91	62
Age	11.6 (0.69)	13.1 (0.75)	14.1 (0.79)	15.4 (0.78)	16.2 (0.86)	17.2 (0.85)	18.2 (0.81)
Male	61.0%	71.3%	52.9%	63.9%	47.0%	62.6%	58.1%
RFL	72.7%	80%	73.1%	80.7%	71.0%	71.4%	72.6%
African American	14.3%	18.8%	32.7%	20.5%	20.0%	17.6%	14.5%
White	23.4%	18.8%	11.5%	12.1%	18.0%	29.7%	25.8%
Hispanic	55.8%	60.0%	51.9%	66.3%	58.0%	50.6%	56.5%
Other	6.5%	2.4%	3.9%	1.1%	4.0%	2.1%	3.2%
TAKS SS	593.03 (76.8)	585.4 (59.9)	649.6 (54.6)	652.8 (71.7)	2016.8 (245.0)	2071.4 (147.3)	2051.0 (197.5)
TAKS LS	751.2 (105.3)	727.1 (97.7)	871.3 (90.7)	886.7 (94.5)	945.3 (166.0)	959.6 (114.3)	990.1 (112.9)
Adequate Comprehender Group							
<i>n</i>	79	100	91	122	128	127	108
Age	11.7 (0.58)	12.9 (0.76)	14.1 (0.72)	15.0 (0.73)	16.1 (0.79)	17.1 (0.74)	18.1 (0.71)
Male	48.1%	47.0%	42.9%	45.9%	43.8%	46.5%	43.5%
RFL	65.8%	66.0%	64.8%	51.6%	58.6%	57.5%	50.9%
African American	12.7%	24.0%	20.9%	21.3%	21.9%	28.4%	18.5%
White	30.4%	19.0%	26.4%	23.0%	30.5%	23.6%	35.2%
Hispanic	55.7%	52.0%	47.3%	50.8%	45.3%	40.2%	41.7%
Other	1.2%	5.0%	5.4%	4.9%	2.3%	7.8%	4.6%
TAKS SS	749.0 (70.5)	752.9 (78.2)	784.1 (66.3)	822.8 (75.9)	2289.6 (93.5)	2340.6 (127.9)	2321.2 (103.9)
TAKS LS	1082.4 (141.5)	1091.8 (157.9)	1161.5 (121.6)	1249.1 (151.4)	1310.1 (127.5)	1310.4 (112.3)	1339.4 (108.5)

Note. *n* = 1352. RFL: Reduced or Free Lunch Status. TAKS SS: TAKS Reading Scale Score. TAKS LS: TAKS Reading Lexile Score.

Table 2

Example Passage from Bridging Inferences Task

Passage: Near textual distance	Passage: Far Textual Distance
Alan sits in the back row of his fourth grade classroom.	Alan does not like getting in trouble with his teacher.
He sits beside two other boys who tell very funny jokes.	Alan sits in the back row of his fourth grade classroom.
Alan heard one of them tell a very funny joke.	He sits beside two other boys who tell very funny jokes.
The two boys giggled.	Alan heard one of them tell a very funny joke.
Alan does not like getting in trouble with his teacher.	The two boys giggled.
Continuation: Good	Continuation: Good
Alan kept quiet.	Alan kept quiet.
Continuation: Poor	Continuation: Poor
Alan laughed loudly.	Alan laughed loudly.

Note. n = 1203.

Table 3

Demographics by Group

	Struggling Comprehenders <i>n</i> = 531	Adequate Comprehenders <i>n</i> = 672
Age	15.1 (2.1)	15.2 (2.1)
% Male	60%	45%
% Free or Reduced Lunch	74%	59%
% African American	21%	22%
% White	19%	26%
% Hispanic	56%	47%
% Other	4%	5%
TAKS Reading Scale Score	1221.7 (716.5)	1508.0 (773.2)
TAKS Reading Lexile Score	877.8 (144.4)	1235.9 (161.3)
Grade 6	11%	9%
Grade 7	14%	14%
Grade 8	19%	13%
Grade 9	15%	17%
Grade 10	17%	17%
Grade 11	15%	16%
Grade 12	10%	13%

Note. *n* = 1203. Struggling Comprehenders: Students attaining scale scores between 1200–2150 on the TAKS Reading Test, first administration. Proficient Comprehenders: Students attaining scale scores above 2150 on the TAKS Reading Test, first administration.

Table 4

Means and Standard Deviations by Reader Subgroup

Measure	Struggling Comprehenders <i>n</i> = 531	Adequate Comprehenders <i>n</i> = 672
Woodcock Johnson-III, Letter Word Identification Scale Score	96.4 (6.9)	102.3 (9.2)
Test of Word Reading Efficiency, Composite Scale Score	90.5 (9.5)	96.2 (10.1)
Test of Sentence Reading Efficiency and Comprehension Scale Score	86.5 (9.6)	97.5 (11.0)
Gates MacGinitie Reading Tests, Scale Score	86.7 (9.9)	97.4 (13.2)
Gates MacGinitie Reading Tests, Lexile Score	826.5 (149.7)	964.1 (176.0)
Kauffman Brief Intelligence Test-2, Verbal Knowledge Scale Score	86.6 (9.3)	99.3 (11.3)
Woodcock Johnson-III, Numbers Reversed Scale Score	90.2 (12.0)	96.7 (13.7)

Note. *n* = 1203. Struggling Comprehenders: Students attaining scale scores between 1200–2150 on the TAKS Reading Test, first administration. Proficient Comprehenders: Students attaining scale scores above 2150 on the TAKS Reading Test, first administration. The Woodcock-Johnson III, Test of Word Reading Efficiency, Test of Sentence Reading Efficiency and Comprehension, and Kauffman Brief Intelligence Test-2 are norm-referenced tests with a mean of 100 and standard deviation of 15 scale score points. The Gates MacGinitie Reading Tests, Scale Score was calculated from the National Curve Equivalent (NCE) as $(100 + NCE - 50) / 21.06 * 15$.

Average performance on the Bridging Inference Task by textual distance and textual consistency by reader sub-group

Table 5

	Near-Consistent		Near-Inconsistent		Far-Consistent		Far-Inconsistent	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Reader Sub-group								
Adequate Comprehenders	.82	.20	.84	.18	.68	.20	.58	.25
Struggling Comprehenders	.72	.22	.70	.24	.57	.23	.50	.24

Note. *N* = 1203.

Table 6

Average performance on the Bridging Inference Task by textual distance and grade level

Grade Level	Textual Distance = Near <i>n</i> = 16 items		Textual Distance = Far <i>n</i> = 16 items		Difference between Near and Far Scores	
	Mean	Std.	Mean	Std.	Mean	Std.
Grade 6	0.694	0.18	0.557	0.16	0.137	0.21
Grade 7	0.727	0.19	0.584	0.17	0.143	0.21
Grade 8	0.764	0.17	0.564	0.17	0.20	0.20
Grade 9	0.770	0.18	0.559	0.18	0.21	0.19
Grade 10	0.828	0.13	0.623	0.19	0.205	0.19
Grade 11	0.802	0.16	0.603	0.18	0.202	0.19
Grade 12	0.816	0.15	0.621	0.17	0.196	0.18

Note. *N* = 1203. Near $F(6, 1194) = 13.25, p < .0001$. Far $F(6, 1189) = 4.47, p = .0003$. Near-Far Difference $F(6, 1189) = 3.84, p = .0009$.

Table 7
Hierarchical Regression Analysis Predicting TOSREC Standard Score among Adequate and Struggling Comprehenders

Step and Predictor	Parameter Estimate	Standard Error	t value	Pr > t	Squared Semi-Partial
Intercept	22.20	5.54	4.01	<.0001	
Grade Level	.95	.21	4.51	<.0001	.004
WJ-III Letter Word Identification	.09	.05	1.82	.07	.22
TOWRE	.27	.04	7.77	<.0001	.09
WJ-III Numbers Reversed	.06	.03	2.28	.02	.01
KBIT-2 Verbal Knowledge	1.59	.18	8.99	<.0001	.12
Reader Group Status	4.87	3.39	1.43	.15	.03
Bridge-It-Near	7.04	3.14	2.24	.03	.007
Reader Group Status X Bridge-It-Near	-1.17	4.33	-.27	.79	.00006

Note. $F(8, 603) = 69.63, p < .0001, R^2 = .48$. WJ-III Letter Word Identification standard score used. TOWRE composite standard score used. WJ-III Numbers Reversed standard score used. KBIT-2 Verbal Knowledge standard score used. Group Status: 0 = Struggling Comprehenders; 1 = Adequate Comprehenders. Bridge-It Near accuracy score used (max 16). Reader Group Status X Bridge-It represents the product of Group Status and Bridging Inference Near score.

Table 8
Hierarchical Regression Analysis Predicting Gates MacGinitie Reading Test- Lexile Score among Adequate and Struggling Comprehenders

Step and Predictor	Parameter Estimate	Standard Error	t value	Pr > t	Squared Semi-Partial
Intercept	-288.51	78.05	-3.70	.0002	
Grade Level	47.39	2.95	16.07	<.0001	.23
WJ-III Letter Word Identification	3.24	.71	4.58	<.0001	.17
TOWRE	.75	.49	1.53	.13	.01
WJ-III Numbers Reversed	.98	.38	2.57	.01	.01
KBIT-2 Verbal Knowledge	17.49	2.47	7.08	<.0001	.08
Reader Group Status	-92.31	48.27	-1.91	.06	.01
Bridge-It-Near	136.12	43.94	3.10	.002	.03
Reader Group Status X Bridge-It-Near	159.84	61.49	2.60	.01	.005

Note. $F(8, 614) = 93.33, p < .0001, R^2 = .55$. WJ-III Letter Word Identification standard score used. TOWRE composite standard score used. WJ-III Numbers Reversed standard score used. KBIT-2 Verbal Knowledge standard score used. Reader Group Status: 0 = Struggling Comprehenders; 1 = Adequate Comprehenders. Bridge-It Near accuracy score used (max 16). Group Status X Bridge-It represents the product of Group Status and Bridging Inference Near score.

Table 9
 Hierarchical Regression Analysis Predicting TOSREC Standard Score among Adequate and Struggling Comprehenders

Step and Predictor	Parameter Estimate	Standard Error	t value	Pr > t	Squared Semi-Partial
Intercept	25.26	5.59	4.52	<.0001	
Grade Level	1.02	.21	4.93	<.0001	.003
WJ-III Letter Word Identification	.09	.05	1.87	.06	.22
TOWRE	.27	.04	7.69	<.0001	.09
WJ-III Numbers Reversed	.06	.03	2.28	.02	.01
KBIT-2 Verbal Knowledge	1.62	.18	9.24	<.0001	.12
Reader Group Status	2.17	2.50	.87	.39	.03
Bridge-It-Far	1.39	3.33	.42	.68	.003
Reader Group Status X Bridge-It-Far	3.84	4.18	.92	.36	.0007

Note. $F(8, 603) = 68.21, p < .0001, R^2 = .48$. WJ-III Letter Word Identification standard score used. TOWRE composite standard score used. WJ-III Numbers Reversed standard score used. KBIT-2 Verbal Knowledge standard score used. Group Status: 0 = Struggling Comprehenders; 1 = Adequate Comprehenders. Bridge-It Far accuracy score used (max 16). Reader Group Status X Bridge-It represents the product of Group Status and Bridging Inference Far score.

Table 10
Hierarchical Regression Analysis Predicting Gates MacGinitie Reading Test- Lexile Score among Adequate and Struggling Comprehenders

Step and Predictor	Parameter Estimate	Standard Error	t value	Pr > t	Squared Semi-Partial
Intercept	-286.78	79.05	-3.63	.0003	
Grade Level	49.83	2.91	17.13	<.0001	.23
WJ-III Letter Word Identification	3.43	.72	4.80	<.0001	.17
TOWRE	.91	.50	1.83	.07	.01
WJ-III Numbers Reversed	.95	.39	2.45	.02	.01
KBIT-2 Verbal Knowledge	18.79	2.48	7.59	<.0001	.08
Reader Group Status	-52.80	35.34	-1.49	.14	.01
Bridge-It-Far	60.17	47.19	1.28	.20	.02
Reader Group Status X Bridge-It-Far	156.08	59.23	2.63	.009	.005

Note. $F(8, 614) = 89.85, p < .0001, R^2 = .54$. WJ-III Letter Word Identification standard score used. TOWRE composite standard score used. WJ-III Numbers Reversed standard score used. KBIT-2 Verbal Knowledge standard score used. Reader Group Status: 0 = Struggling Comprehenders; 1 = Adequate Comprehenders. Bridge-It Far accuracy score used (max 16). Group Status X Bridge-It represents the product of Group Status and Bridging Inference Far score.