

# **The effect of altered Lexile levels of the same text on reading comprehension**

**Catherine L. Rand**

## **Abstract:**

The purpose of this study was to examine how using simplified versions of a text will affect reading comprehension. 335 students in grades 4-8 read one of five versions of an informational text retrieved from Newsela and then took a comprehension test. Results from a 3-way ANOVA showed no significant interaction between grade, reading level and text condition. Pairwise comparisons showed that below-level readers' scores differed from the scores of on-level or above-level readers only when given extremely lower levels of texts. Regression analysis showed no significant contribution of text level to overall comprehension scores. This study is important because it raises questions about our decision-making in choosing leveled texts for reading instruction.

The purpose of this study was to examine how using texts that have been simplified will affect the reading comprehension scores of students of different reading abilities in grades four through eight.

It is well established that to comprehend a text meaningfully and most effectively learn from it, students need texts within their zone of proximal development (Chaiklin, 2003). The Lexile Framework for Reading is a widely used formula used to measure text complexity on a discrete scale from 100 to 1200 (Smith, Stenner, Horabin, & Smith, 1989). Recently, Lexile's formula has been used by sites like Newsela.com to not only measure the complexity of a text, but to create simplified versions of informational articles at different levels (Newsela, 2019). In theory, this would allow differentiation by reading ability in a classroom without having to use

different content for different levels of readers. However, there has been some dispute as to the validity of using only quantitative measures to gauge text complexity (Cunningham, Hiebert, & Mesmer, 2018; White & Clement, 2001), the influence of cultural and situational factors on comprehension (Glenberg, 2011), and the effectiveness of altering text to be more or less complex (Lupo, Tortorelli, Invernizzi, Ryoo, & Strong, 2019).

Some of the factors Lexile uses to determine text complexity are the average length of words per passage, average number of words per sentence, and the length of the passage (Smith et al., 1989). Newsela does not release information on how it simplifies the texts. However, with individual inspection of differently-leveled articles, it seems likely that to make texts less complex, they shorten the number of words per text, reduce the length of words, and shorten sentences (Alva-Manchego, Bingel, Paetzold, Scarton, & Specia, 2017; Xu, Callison-Burch, & Napoles, 2015). There has been criticism that by simplifying these texts, the coherence of the text suffers, making it more difficult, not easier, to read (Reed & Kershaw-Herrera, 2016; Cunningham et al., 2018). There is evidence to suggest supporting readers with a higher-level text can be just as effective as assigning them a lower-level text (Lupo, Strong, Smith, 2019; Shanahan, Fisher, & Frey, 2012). Finally, there is a concern about equity in assigning easier texts, as they are typically shorter and introduce students to fewer unfamiliar vocabulary words (Lupo et al., 2019). Amendum, Conradi, and Hiebert (2018) reviewed the research on the effect of differently leveled texts on comprehension and found mixed results. There hasn't been much research done to assess the effect of using different levels of the same article on comprehension, or on Newsela's text simplification process, but some work has been done to create a generalizable understanding of the text simplification process (Alva-Manchego et al., 2017).

In order to examine the effect of Newela's text alterations on student comprehension, we set out to answer the following research questions: Is there an interaction between reading ability, grade level and different Lexile versions of the article on reading comprehension? Is there a difference in reading comprehension scores of students who read different Lexile versions of the article? Is the effect of reading ability on reading comprehension influenced by the Lexile level of the text?

## **Methods**

**Subjects:** We recruited all the students in grades 4, 5, 6, 7, and 8 from two urban schools and included the 335 students who agreed to participate. The racial and ethnic background of students is shown in Table 1.

Students were grouped into three categories: below-level readers, on-level readers, and above level readers based on a combination of the previous year's PARCC scores, benchmark reading tests scores, and teacher identification, as shown in Table 2.

**Materials:** We downloaded the Newsela article, *Fishy looking robot SoFi is helping scientists spy on life under the sea* (Netburn, 2018) in the following five different formats from Newsela.com:

- Lexile Level 560
- Lexile Level 820
- Lexile Level 1060
- Lexile Level 1130
- Lexile Level 1250

Subjects were randomly assigned to a different Lexile Leveled text, as shown in Table 3.

We created a 10-question multiple choice comprehension test based on the Common Core State Standards for Reading Informational Text. All subjects completed the same

comprehension test. Using a multiple linear regression, we found a strong correlation between comprehension score and reading ability level, indicating the validity of the test for measuring comprehension.

**Procedures:** We provided each teacher with a packet of articles that were randomly mixed with equal numbers of each article format. The texts were coded so that neither the teacher nor the students could tell the format of each article. The teacher passed out these texts to the students in no particular order to randomize the assignment of articles. Students read the article and answered ten comprehension questions. They were told the questions would count as a quiz grade. All the students in the study had previously read Newsela articles and answered comprehension questions as classroom assignments.

Data sources included comprehension test scores which ranged from 0-10 as the dependent variable. Independent variables were Text Level (five levels of Lexile versions) of the article, Reading Ability Level (above level, on level, below level) and Grade Level (4, 5, 6, 7, or 8).

A three-way ANOVA was conducted to determine the effects of Reading Ability Level, Grade and Text Level on comprehension score. We then performed pairwise comparisons using a Bonferroni correction. A multiple linear regression was run to determine how much of the variation in comprehension scores was determined by the independent variables.

## **Results**

There was no statistically significant three-way interaction between Reading Ability Level, Grade, and Text Level on the 3-way ANOVA,  $F(24, 180) = .971, p = .507$ . The three-way ANOVA showed that the group means were not significantly different and, therefore, there was

no evidence in this sample to show that students who read different Lexile versions of the article had significantly different comprehension scores.

Next, all simple pairwise comparisons were run for Comprehension Score with a Bonferroni adjustment applied. There was a statistically significant simple two-way interaction between Reading Ability Level and Text Level Condition for below level readers,  $F(4, 264) = 3.649, p < .007$  but not for on-level readers,  $F(4, 264) = .876, p = .479$  or above-level readers,  $F(4, 264) = .109, p = .979$ . See Table 4 that shows the statistically significant pairwise comparisons of Reading Level and Test Condition on Comprehension Score.

When we looked at only the difference between reading ability and text level, we found a significant difference in comprehension scores of below level readers between the those who read articles at the 560 and those who read articles at the 1130 Lexile level ( $p < .008$ ) and between those who read 820 and those who read 1130 Lexile levels ( $p < .024$ ). There was no difference in scores of below level readers who read articles at closer Lexile levels, for example 530 compared to 820 or 1060. The difference was only significant when there was a larger variation in Lexile range.

To determine how much of the variation in the comprehension scores was explained by the independent variables, we conducted a multiple linear regression.  $R^2$  for the overall model was .211% with an adjusted  $R^2$  of .204%, a small size effect according to Cohen (1988). The three independent variables combined, Reading Ability Level, Grade Level and Text Level, statistically significantly predicted reading comprehension scores,  $F(3, 330) = 29.459, p < .001$ . However, Text Level condition was not a statistically significant predictor of Comprehension Test Score. The slope coefficient was  $-.109$  showing that for every 1 level decrease in the Text

Level (which ranged from 70-250 Lexile points), the comprehension test score increased by .109 points on a 10-point scale.

We also used descriptive statistics to analyze student performance on each question by calculating the percentage of participants who correctly answered the question at any given Lexile level of text. See Figure 1 for the full results. On none of the ten questions was there a linear reduction in the number of correct answers per the rise in Lexile level of the text. Similarly, there was no discernable pattern to which Lexile level of text helped students perform better across questions. On question four, which measured main idea, the lowest level of text showed higher rates of correct answers. On questions three and ten, however, which measured key details and author's purpose respectively, the highest Lexile version was tied for the highest rate of correct answers. Question three's other highest percentage correct was the second-to-lowest Lexile level. Question ten's other highest percentage correct was the lowest Lexile level of the text.

The comprehension questions were administered in an order that mimics the CCSS standards, with the more basic tasks of finding key details and main idea asked before the more challenging standards like comparing ideas or the evaluation of the author's use of evidence. See Table 5 for the list of questions and their alignment to CCSS. Even considering the complexity of the task asked by the question, there was no explainable change in participant performance. The highest two scoring questions by all participants were question two, that measured inference, and question ten, that measured author's purpose. Questions six and nine, that measure text structure and illustration inference respectively, were the questions that participants scored the lowest on, regardless of the level of text they received.

In summary, our results showed the following:

- Above-level readers performed at a higher level than on-level readers or below-level readers at all Lexile levels
- Using a lower Lexile article only affected the scores of below-level readers, not on-level or above-level readers
- The effect of altering test conditions only affected comprehension scores by an average of .1 on a scale of 0-10.
- Though there was considerable difference in performance between questions, there was not a relationship found between the level of text and participants' performance in any of the ten questions

## **Discussion**

This study is important because it adds to our decision-making ability in choosing texts for reading instruction and practice. Teachers are currently using reading texts, including Newsela, that have been simplified to achieve different readability levels (Amendum, Conradi, & Hiebert, 2018). We have assumed that having texts at the appropriate Lexile level would improve comprehension for all readers, yet this study calls into question that assumption. There was some evidence that for below-level readers, having a lower level text improved their comprehension scores, however, this effect was very small. For on-level and above-level readers at all grade levels from 4-8, there was no significant difference in comprehension test scores across the five Lexile levels of the different texts.

Researchers have previously expressed concern that the process of simplifying texts in order to lower the readability might reduce some of the linguistic information that helps comprehension (Reed & Kershaw-Herrera, 2015; Xu, Callison-Burch, & Naples, 2015). We did

not find that the simplified texts interfered with comprehension and helped the below-level readers when there was a wide range of Lexile difference.

This research was done as an exploratory study and has limitations. We used only one text set, so there could be anomalies in this particular manipulation of text. We also used only two schools that may not be representative of other learners. Moving forward, we need to examine other types of simplified texts with more student populations to be able to generalize any findings. We expect to have additional data on other text sets and subjects for our final paper.

This research supports the theme of collaboration with organizational stakeholders for this year's annual meeting in that it was a university-partnership collaboration. The first author is a classroom teacher who worked with University researchers to collect and analyze this research. By collaborating on this inquiry, we cross boundaries and ensure that the research findings will be useful to the K-12 community by offering insight into popular instructional materials.

## References

- Alva-Manchego, F., Bingel, J., Paetzold, G., Scarton, C., & Specia, L. (2017, November). Learning how to simplify from explicit labeling of complex-simplified text pairs. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 295-305).
- Amendum, S.J., Conradi, K., & Hiebert, E. H. (2018). Does text complexity matter in the elementary grades? A research synthesis of text difficulty and elementary students' reading fluency and comprehension. *Educational Psychology Review*, 30(1), 121-151.
- Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. *Vygotsky's educational theory in cultural context*, 1, 39-64.
- Cohen, Jacob. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, N.J: L. Erlbaum Associates; 1988.

- Cunningham, J. W., Hiebert, E. H., & Mesmer, H. A. (2018). Investigating the validity of two widely used quantitative text tools. *Reading and Writing*, 31(4), 813-833.
- Glenberg, A. M. (2011). How reading comprehension is embodied and why it matters. *International Electronic Journal of Elementary Education*, 4(1), 5-18.
- Lupo, S. M., Tortorelli, L., Invernizzi, M., Ryoo, J. H., & Strong, J. Z. (2019). An exploration of text difficulty and knowledge support on adolescents' comprehension. *Reading Research Quarterly*.
- Lupo, S. M., Strong, J. Z., & Conradi Smith, K. (2019). Struggle is not a bad word: Misconceptions and recommendations about readers struggling with difficult texts. *Journal of Adolescent & Adult Literacy*, 62(5), 551-560.
- Netburn, D. (2018). *Fishy looking robot SoFi is helping scientists spy on life under the sea*. Retrieved on June 18, 2018 from <https://newsela.com/read/robotic-fish-ocean-spy/id/41688/>
- Newsela.com. (2019) <https://support.newsela.com/hc/en-us/articles/360008500391-Grade-to-Lexile-Conversion>
- Reed, D. K., & Kershaw-Herrera, S. (2016). An examination of text complexity as characterized by readability and cohesion. *The Journal of Experimental Education*, 84(1), 75-97.
- Shanahan, T., Fisher, D., & Frey, N. (2012). The challenge of challenging text. *Educational leadership*, 69(6), 58–62. Retrieved from <https://search-ebSCOhost-com.proxy.libraries.rutgers.edu/login.aspx?direct=true&db=aph&AN=73183264&site=ehost-live>
- Smith, D. R., Stenner, A. J., Horabin, I., & Smith, M. (1989). *The Lexile scale in theory and practice: Final report for NIH grant ID-19448*. New Orleans: International Reading Association.
- White, S., & Clement, J. (2001). *Assessing the Lexile Framework: Results of a panel meeting*. US Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.
- Xu, W., Callison-Burch, C., & Napoles, C. (2015). Problems in current text simplification research: New data can help. *Transactions of the Association for Computational Linguistics*, 3, 283-297.

Percentage of Students	White	Black	Hispanic	Asian	2 or More Races	Free & Reduced
School A	27.4	15.5	20	27.4	7.6	37%
School B	25.9	10.7	28.9	28.3	5.9	38%

Table 1: Ethnic/racial background of subjects

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Below Level Reading Ability	15	8	23	16	11
On-Level Reading Ability	30	33	38	27	45
Above Level Reading Ability	13	18	12	15	31
<b>Total:</b>	58	59	73	58	87

Table 2: Number of subjects at each reading level and each grade.

	Lexile Level 560 (Level 1)	Lexile Level 820 (Level 2)	Lexile Level 1060 (Level 3)	Lexile Level 1130 (Level 4)	Lexile Level 1250 (Level 5)
Grade 4	13	13	12	8	12
Grade 5	12	12	11	13	11
Grade 6	15	15	17	13	13
Grade 7	10	12	14	11	11
Grade 8	16	15	17	20	19
<b>Total</b>	66	67	71	65	66

Table 3: Number of subjects assigned in each grade to each Lexile level

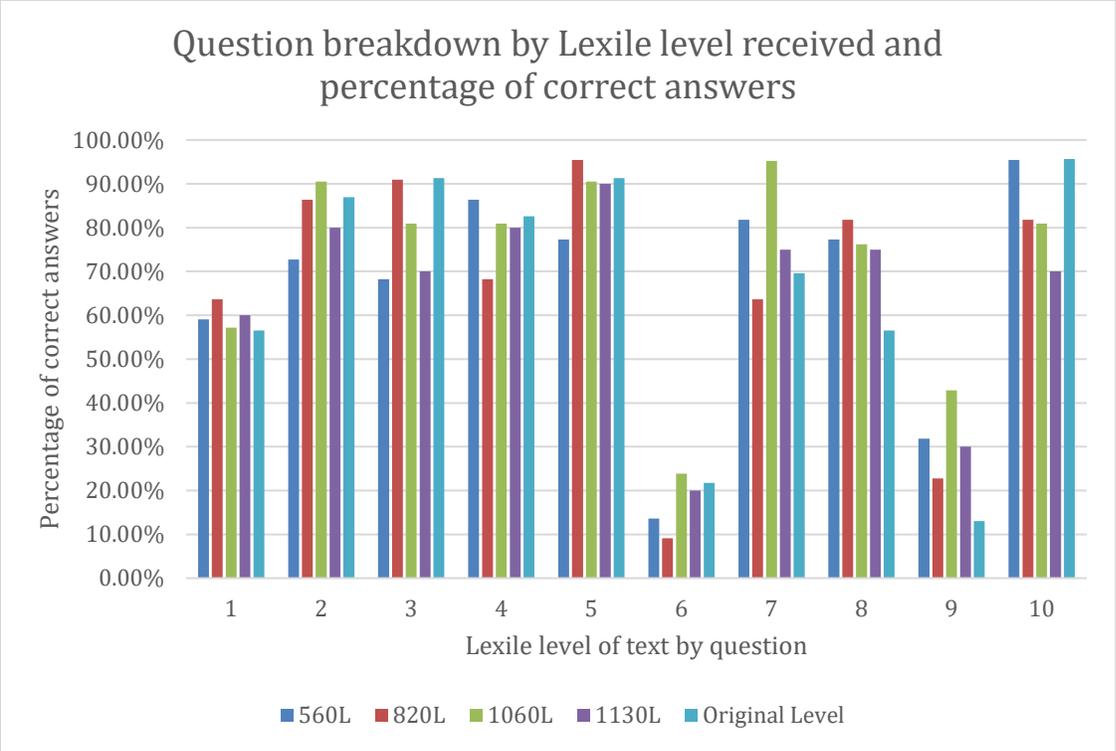
Reading Ability Level	Pairwise Comparison with Text Level Confiton	Sig. with Bonferroni Adjustment
Below Level	Text Level 560 *Text Level 1130	.008
	Text Level 820*1130	.024

<b>Text Level Condition</b>	<b>Pairwise Comparison with Reading Ability Level</b>	<b>Sig. with Bonferroni Adjustment</b>
Text Level 820	Below Level * Above Level	.041
Text Level 1060	Below Level * Above Level	.008
Text Level 1130	Below Level* On Level	<.001
	Below Level * Above Level	<.001
Text Level 1250	Below Level * On Level	.002
	Below Level *Above Level	.001

*Table 4: Statistically significant pairwise comparisons of Reading Level and Test Condition on Comprehension Score*

<b>CCSS Standard</b>	<b>Aspect of comprehension measured</b>	<b>Comprehension question administered</b>
RI.1	Key Details	Question 1: What makes SoFi the first of its kind? Question 2: Why does the scientist say the fish is “magical”? Question 3: How do the wild fish react to SoFi?
RI.2	Main Idea	Question 4: What is the central idea of the article? Question 5: How do scientists want to improve SoFi?
RI.3	Relationship between ideas	Question 7: How does the beginning and end of the article show SoFi differently?
RI.5	Text Structure	Question 6: Which of the following text structures is most represented by this article?
RI.6	Author’s Purpose	Question 10: What is most likely the reason the author wrote this article?
RI.7	Media Literacy	Question 9: What point in the article is most demonstrated by the illustrations?
RI.8	Reason and Evidence	Question 8: What evidence shows why the scientists released SoFi in Fiji in the South Pacific Ocean?

*Table 5: Content description of comprehension questions*



*Figure 1: Question breakdown by Lexile level received and percentage of correct answers*