THE ASSOCIATIONS BETWEEN READABILITY MEASURES AND PROBLEM SOLVING IN ALGEBRA

Candace Walkington
Southern Methodist University
cwalkington@smu.edu

Virginia Clinton
Univ. of Wisconsin-Madison
vclinton@wisc.edu

Elizabeth Howell
Southern Methodist University
ehowell@smu.edu

Solving mathematics story problems requires text comprehension skills. However, previous studies have found few connections between traditional measures of readability and performance on story problems. We hypothesized that recently-developed measures of readability may illuminate associations between text difficulty and problem solving. We used data from 3,394 middle and high school students solving algebra story problems in Cognitive Tutor Algebra. We found that several indicators of readability were negatively associated with incorrect answers. Moreover, indicators of readability were associated with students requesting hints from Cognitive Tutor. These findings are discussed in the context of models of algebra problem solving and previous research on text difficulty in solving story problems.

Keywords: Curriculum, Cognition, Problem Solving

Theoretical Framework

Algebra word problems, or story problems, can provide access to important mathematical concepts. Their concrete, verbal language can connect to students' prior experiences, providing a conceptual foundation on which more advanced skills can be built (Koedinger & Nathan, 2004; Walkington, Sherman, & Petrosino, 2012). However, in order to understand and solve story problems, students must be able to comprehend and interpret the story text (Kintsch & Greeno, 1985). This idea is articulated in Nathan, Kintsch, and Young's (1992) framework for algebra story problem solving, where students develop a mental representation of the actions and relationships in a story problem's text, a *situation model*, as well as a mental representation of the formal mathematics, a *problem model*. The coordination of situation and problem models is a key element of learning from algebra story problems (Walkington et al., 2012).

Based on this framework, it is reasonable to assume that the readability of story problems contributes to their difficulty. However, researchers have found minimal evidence linking the readability with problem solving accuracy (Wiest, 2003). One reason may be the use of traditional measures of readability that assess the text based on the familiarity of the words, the number of syllables, and the number of words per sentence (Graesser, McNamara, & Louwerse, 2012). These measures provide coarse estimates of a text's difficulty, but do not, for instance, capture the *cohesion* of text – the "linguistic glue" that joins the events and ideas, helping readers to understand connections and relationships (Gernsbacher, 1997; McNamara, Graesser, McCarthy, & Cai, in press). For example, the use of consistent verb tense (e.g., past or present) in different sentences can cue the reader that the actions in these sentences occurred during the same time period. Research in elementary mathematics has also suggested that story problems with clear and explicit action (e.g., joining, separating) facilitate problem solving (Carpenter et al., 1999). Given Nathan et al.'s (1992) framework, measures relating to cohesion and explicit action may be a critical component of solving story problems.

In the present study, we examine the texts of a large set of algebra story problems solved by students in a variety of school settings. We argue that some types of story texts may better facilitate the construction of meaningful situation models and their coordination with problem models. We have two research questions: (1) How are different measures of text readability

associated with students' accuracy when solving algebra story problems? and (2) How are different measures of text readability associated with students' help-seeking when solving algebra story problems? By investigating these questions, we seek to better understand the previously unclear relationship between text difficulty and solving story problems.

Method

Cognitive Tutor Algebra (CTA) is a computer-based intelligent tutoring system for Algebra I in use in schools across the United States. The tutoring system presents students with algebra problems and provides adaptive hints if the student requests them (Morgan & Ritter, 2002). As students work through the program, detailed log files are created that store students' interactions with the system. We used the CTA log files from 9 high schools and 1 middle school containing N = 3394 students with active CTA accounts. These schools were in 10 different states, with 2 rural, 4 urban, and 4 suburban schools. Three schools had 0-33% of students eligible for free/reduced lunch, four had 33-66% eligible, two had 66-100% eligible, and one did not have this information available. Five schools had student populations that were predominately White students, three were predominantly African-American students, and two were predominately Hispanic students. Schools also varied in achievement on standardized mathematics assessments.

Data were collected from the first 8 units in CTA that used story problems with linear functions; CTA is adaptive to student needs, so not all students receive all problems. After demo problems, problems with missing data, and problems with data from fewer than 15 students were omitted, 282 problems remained. On average, each problem had been solved by 398 students (SD = 321). Each story problem had several parts and required students to write a symbolic expression stating the linear relationship(s) in the story and fill out a table; some problems also had students construct graphs. See carnegie.learning.com for examples of problems.

In the present study, students' average error rate on each problem and their rate of asking the system for a hint were compiled from the CTA log files. Then the text of the introduction to the story problem was entered into the Coh-Metrix 3.0 software. The introduction can be one or more sentences, and generally describes the slope and intercept terms of the story (see Table 2). Coh-Metrix is a software tool developed in 2002 that provides multiple, varied, and precise measures of readability (e.g., frequency of different types of words like pronouns and verbs), measures of cohesion (e.g., degree of narrative language or time-related cues), in addition to traditional measures of readability (e.g., number of sentences or word syllables; McNamara et al., in press). Coh-Metrix measures were entered into regression models to test if they significantly predicted the percentage of attempts on that problem that were incorrect, or the percentage of attempts in which hints were requested. The regression models controlled for aspects of the story's mathematical structure, including the structure of the linear function(s) (e.g., positive slope/no intercept, negative slope/negative intercept), the type of numbers used (e.g., fractions, large whole numbers), and the unit and section the story problem came from.

Results and Discussion

Accuracy

Table 1 shows only the readability measures from Coh-Metrix that significantly predicted students giving inaccurate responses: Text Easability PC Temporality, and incidence of adverbs, intentional verbs and third person singular pronouns. For the scale on which each of these values was measured by Coh-Metrix (to interpret the B values), see http://cohmetrix.memphis.edu. "Text Easability PC Temporality" measures how many cues about temporality the text contains, and how consistent tense (e.g., present or past) and aspect (perfect or imperfect) are. This captures how well sequences of events that occur in the past, present, or future are made explicit,

including cues like "an hour later." Consistent verb tense provides coherence that helps the student connect the information in different sentences (Gernsbacher, 1996). Thus it is not surprising that Table 1 shows that as temporal cohesion increases, incorrect answers decrease (indicated by the negative B value). Related to this is the result in Table 1 that as incidence of adverbs increase, incorrect answers decrease. Adverbs can provide temporal cues, which may help the reader follow the time course of the story. Adverbs can also add descriptive information about how and to what extent actions and events occurred. This information may enrich the story context, providing cues that support situation model construction. Examples of problems with both high and low scores on these measures can be seen below in Table 2.

Table 1: Regression results for the impact of readability on problem solving accuracy

	B	**SE(B)**	***T***	***p***
(Intercept)	21.153	2.096	10.094	0.0001
Text Easability PC Temporality	-0.030	0.008	-4.042	0.0002
Adverb incidence	-0.045	0.011	-4.176	0.0004
Intentional verbs incidence	-0.026	0.011	-2.290	0.0154
Third person singular pronoun incidence	-0.024	0.011	-2.182	0.0284

Table 2: Sample problems for comparison of readability scores in two measures

Measure	**High**	**Low**
Text Easability PC Temporality	We ***work*** for a company that ***makes*** and ***sells*** hot-air balloons. Each hot-air balloon ***sells*** for $3500. Our warehouse ***has*** 20 balloons that ***have already been*** sold but ***have not yet been*** shipped. (Coh-Metrix 98.26 percentile)	We ***are building*** doghouses and ***have already finished*** five of them. A local store ***agrees to buy*** as many as we ***can make*** for $25 apiece. (Coh-Metrix 0.03 percentile)
Adverb Incidence	Many people do not realize that when one issues a false fire alarm the community ***needlessly*** spends thousands of dollars to protect the citizens. The money that gets wasted and ***eventually*** translates into higher taxes for everyone is approximately $4000 per call. This includes the costs to maintain the trucks and equipment the ***very*** expensive insurance and the salaries of the firefighters and other employees. This financial loss to the community is the reason why the fines for ***such*** illegal activity are ***so*** high. (Coh-Metrix score 107.143).	A friend offers us a job selling snow cones. Since he owns the snow cone maker he will let us use it for $50. We will make $0.50 on each snow cone we sell. (Coh-Metrix score 0)

Table 1 also shows that as intentional verbs and third person pronouns in a story problem increase, incorrect answers decrease. Story problems are easier for students to solve when there are purpose-driven, goal-oriented actions like buying a house or cooking a meal. The concrete information provided by intentional verbs may assist the student in developing a more vivid and meaningful situation model. Pronouns also act as facilitators when they have clear referents in the story text (White, 2012). Pronouns provide cues that prompt the student to connect the new information containing the pronoun with the old information containing the referent.

Help-Seeking

Table 3 shows the readability measures from Coh-Metrix that significantly predicted students requesting a hint from Cognitive Tutor. Story problems that had more sentences in them were more likely to promote students asking for hints, rather than attempting the problem; a long story text may be discouraging and increase perceived difficulty (White, 2012). Temporal connectives like "first" and "then" were associated with more hint-seeking, which is unexpected given other temporal coherence cues were negatively associated with inaccurate answers. Students who see a text with multiple temporal cues may be intimidated by the amount of information to process.

These students may seek a hint even though the temporal connectives improve situation model construction. Finally, we see that stories with third person singular pronouns are associated with less hint-seeking, which is consistent with the problem-solving accuracy findings.

Table 3: Regression results for the impact of readability on hint requests

	B	SE(B)	*t*	*p*
(Intercept)	2.398	0.868	2.762	0.0018
Sentence count, number of sentences	0.343	0.099	3.470	0.0004
Temporal connectives incidence	0.013	0.005	2.843	0.0042
Third person singular pronoun incidence	-0.014	0.004	-3.282	0.0012

Significance

The purpose of this study was to examine associations between readability measures and solving algebra story problems, using recent advances in software-based text mining. Our results suggest important implications for mathematics teaching and curriculum design. Formulating problems in which the student is cued to connect related events appears to facilitate problem-solving. Avoiding overly lengthy story contexts and overuse of temporal cues may also provide initial access. Future studies could vary the readability levels of problems on these indicators, and examine student learning. These results are different from previous findings on this topic in which traditional, coarse measures of readability have been used. According to the model we presented for algebra story problem-solving (Nathan et al., 1992), our findings suggest that students were better able to solve a story problem if its text supported their construction of a situation model and its coordination with their problem model.

References

Carpenter, T., Fennema, E., Franke, M., Levi, L., & Empson, S. (1999). *Children's Mathematics: Cognitively Guided Instruction.* Heinemann.

Gernsbacher, M. A. (1996). Coherence cues mapping during comprehension. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships in the production and comprehension of text* (pp. 3-21). Mahwah, NJ: Erlbaum.

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix Providing Multilevel Analyses of Text Characteristics. *Educational Researcher*, *40*(5), 223-234.

Graesser, A.C., McNamara, D.S., & Louwerse, M. (2012). Sources of text difficulty: Across the ages and genres. In J.P. Sabatini & E. Albro (Eds.), *Assessing reading in the 21st century: Aligning and applying advances in the reading and measurement sciences.* Lanham, MD: R&L Education.

Kintsch, W. & Greeno, J.G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*(1), 109-129.

Koedinger, K., & Nathan, M. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences*, *13*(2), 129-164.

McNamara, D.S., Graesser, A.C., McCarthy, P., & Cai, Z. (in press). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge: Cambridge University Press.

Morgan, P., & Ritter, S. (2002). An experimental study of the effects of Cognitive Tutor Algebra I on student knowledge and attitude. Pittsburgh, PA: Carnegie Learning, Inc.

Nathan, M., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction*, *9*(4), 329-389.

Walkington, C., Sherman, M., & Petrosino, A. (2012). 'Playing the game' of story problems: Coordinating situation-based reasoning with algebraic representation. *Journal of Mathematical Behavior*, *31*(2), 174-195.

White, S. (2012). Mining the Text: 34 Text Features That Can Ease or Obstruct Text Comprehension and Use. Literacy Research and Instruction, 51(2), 143-164.

Wiest, L. (2003). Comprehension of mathematical text. *Philosophy of mathematics education journal, 17*, 458.