

Cohesion and Structural Organization in High School Texts

Erin J. Lightman, Philip M. McCarthy, David F. Dufty, and Danielle S. McNamara

University of Memphis
elightman@mail.psyc.memphis.edu
pmccarthy@mail.psyc.memphis.edu
d.dufty@mail.psyc.memphis.edu
d.mcnamara@mail.psyc.memphis.edu

Abstract

Recent research in reading comprehension supports the hypothesis that readers are aided by textual cohesion. Traditional readability formulas are not able to effectively assess levels of textual cohesion, nor do they account for potential comprehension obstacles caused by differences in genre. This research employs the computational tool, Coh-Metrix, to assess distributions of both cohesion and difficulty within chapters of expository and narrative texts. We sampled representative sections from the beginning, the middle, and the end of each chapter in three science textbooks, four history textbooks, and six narrative texts. The results suggest that the three domains differ significantly in terms of structural organization and levels of difficulty and cohesion. Differences between these results are discussed.

Introduction

Selecting texts for students that are appropriate in terms of difficulty is one of the many challenges that teachers face today (Bowen, 1999; Snow, 2002). This is a particularly important problem for demanding texts (such as science), which are not always easy to comprehend (Beck, McKeown, Sinatra, & Loxterman, 1991; McNamara, 2001; McNamara & Kintsch, 1996). It is reasonable to assume that a teacher's decision is based in part on the grade-level assigned to the text by the publisher. This can be problematic as traditional readability formulas usually play a large role in determining a textbook's grade level, leaving teachers and textbook writers to presume that similar difficulty levels exist across genres. The many differences between genres, in addition to recent understanding of text comprehension, suggest that the whole approach of assessing texts with shallow difficulty metrics needs to be reevaluated (Graesser, McNamara, Louwerse, & Cai, 2004).

Recent research has emphasized the importance of cohesion in text comprehension (e.g., Britton & Gulgoz, 1991; McNamara, 2001). Cohesion is the degree to which ideas in a text are explicitly related to each other, thereby facilitating a unified situation model for the reader. However, while cohesion can facilitate learning, more cohesion is not necessarily always better. McNamara, Kintsch, Songer, and Kintsch (1996) found that high cohesion texts predominately aid comprehension for low-knowledge students. This is because there are fewer conceptual and structural gaps in a high cohesion text, thus alleviating the cognitive burden of generating inferences (Graesser, 1993; Linderholm et al., 2000; McNamara, 2001).

Conversely, texts with lower cohesion can aid high-knowledge, highly skilled readers by pushing them to generate more inferences which subsequently increases the probability of knowledge retention (McNamara, 2001). The greater amount of knowledge available to high-knowledge readers means that the inferences they generate have a better chance of being successful, and consequently, high-knowledge readers can benefit from low-cohesion text, especially when the information in the text and their background knowledge become more integrated as a result of generating more inferences. O'Reilly and McNamara (in press) examined these interactions between reading skill, prior knowledge and textual cohesion. They found that the low-knowledge readers benefited from high cohesion text and that reading skill detected whether high-knowledge readers benefited from high-cohesion text.

We extrapolate from these findings a hypothesis that the optimal organization of an expository text would fall into one of two patterns: First, if we assume that students may learn little from a text as they read, then a text chapter should remain consistently high in cohesion. In contrast, if we assume that students gain knowledge as the text progresses and if the topics in the text are not varying widely, then a chapter should gradually fall in cohesion as it progresses. A poorly organized chapter, on the other hand, would exhibit a gradual rise in cohesion. Such a rise would indicate that students typically begin a unit with

higher knowledge than they end it. Because the majority of research on cohesion's role in comprehension and recall has addressed expository texts as opposed to narratives, we suggest that our hypothesis of optimal cohesion distribution would not necessarily apply to narrative texts.

In this paper, we used chapters from current school textbooks and works of literature commonly studied in high school English classes to test our hypothesis of optimal cohesion distribution.

Concerns with Current Metrics

Over the years, numerous metrics have been created to evaluate the difficulty of texts. Readability formulas, such as the Flesch-Kincaid Grade Level (FKGL, Klare, 1974-1975), have been criticized due to the limitations posed by their strong focus on shallow textual features, such as sentence length and word length (Graesser et al., 2004). Indeed, shallow textual features have considerable validity for assessing difficulty, as longer words tend to be less frequently used (Zipf, 1949), and less frequently used words tend to be processed more slowly (Just & Carpenter, 1980). However, these shallow aspects alone explain only a part of text comprehension, ignoring many language and discourse features that are theoretically influential at determining comprehension difficulty (Graesser et al., 2004; McNamara, 2001).

Another problem with traditional metrics that evaluate difficulty based on sentence and word length is their vulnerability to manipulation. As Graesser et al. (2004) discuss, it is a somewhat common practice for textbook writers to adjust the grade level of their course books by simply shortening sentences, which in turn, generally lowers cohesion.

Shallow metrics also fail to recognize that different genres present different problems to their readership. Science texts tend to present unfamiliar topics, meaning that a lower knowledge base can be assumed and that greater cohesion would be beneficial to readers. History and narrative texts, however, tend to be presented more as a chronological series of events on topics with which the reader is more likely to be familiar (Tonjes, Ray, and Zintz, 1999).

Because of their different structures and intentions, traditional and even new metrics, may be able to effectively illustrate features for some genres better than others. For example, a high difficulty level for a science text may indicate characteristics completely different from the same high difficulty level in a narrative. In one of the few empirical studies that have investigated textual distinctions between genres, McCarthy, Graesser, and McNamara (2006) were able to use Coh-Metrix to distinguish between science and narrative texts. Their results suggested that expository texts are more inter-textually cohesive, and narrative texts incorporate more inter-clausal and temporal cohesion markers.

In brief, current readability metrics do not fully capture potential comprehension. As a result, textbook writers cannot be fully confident that their texts are optimally designed for comprehension, and teachers cannot be sure that they are teaching their students with materials of appropriate difficulty. A better approach may be to provide a tool that measures text in terms of both difficulty and cohesion, thereby providing textbook writers and teachers with an alternative assessment of a text's suitability to its intended audience.

Assessing Texts with Coh-Metrix

Coh-Metrix is a computational tool for approximating over 400 indices of textual cohesion, difficulty and conceptual characteristics (Graesser et al., 2004). The system evaluates texts using lexicons, pattern classifiers, part-of-speech taggers, syntactic parsers, shallow semantic interpretations, and other components developed in computational linguistics research (Jurafsky & Martin, 2000; McNamara, Louwerse, & Graesser, 2006). Coh-Metrix incorporates a large number of lexicons including Celex (Baayen, Piepenbrock, & Van Rijn, 1993), WordNet (Miller, Beckwith, Feldbaum, Gross, & Miller, 1990), the MRC Psycholinguistic Database (Coltheart, 1981), and lexicons that estimate the frequency of word use. These lexicons allow Coh-Metrix to assess each word on number of syllables, abstractness, imagery, ambiguity, frequency of usage, age of acquisition, to name just a few (McNamara, Louwerse, & Graesser, 2006). From the output, Coh-Metrix offers an assessment of the cohesiveness and readability of practically any given text.

Coh-Metrix has been used to demonstrate the importance of cohesion in various studies which have measured differences in students' recall of a text based on their prior knowledge, reading skill, and the cohesion of the text. Ozuru, Dempsey, Sayroo, and McNamara (2005), for example, used Coh-Metrix to rate high and low cohesion versions of biology texts and found that subjects benefited most from high cohesion versions. Best, Floyd, and McNamara (2005) investigated 61 third graders' reading comprehension by using the tool to evaluate their recall of narrative and expository texts. The study suggested that children with low levels of world knowledge were more prone to comprehension problems with expository texts.

Coh-Metrix has also been used to examine cohesion and conceptual characteristics in large corpora of texts. McCarthy, Briner, Rus, and McNamara (in press) investigated whether structural cohesion could be analyzed using Coh-Metrix LSA indices. And Louwerse, McCarthy, McNamara, and Graesser (2004) used Coh-Metrix to investigate variations in cohesion across written and spoken texts, finding evidence for the significant distinction between these modes. In this study, we build on previous research and use Coh-Metrix to assess how well a

sample of current student texts correlates with our hypothesis of optimal cohesion distribution as compared to a traditional metric.

Predictions

A textbook rarely covers a single topic. Instead it is divided into many chapters, with each chapter representing a new area of knowledge to be taught and learned, so we consider each chapter of a book to be a self-contained unit. We also consider that an author of a text unit would make the following three assumptions: that the reader is low in knowledge at the beginning of the chapter, that the reader is reading the text to become higher in knowledge, and that the reader should end that unit with higher knowledge or, at least, not with lower knowledge. With these assumptions in mind, our hypothesis of optimal distribution of cohesion would suggest one of two possible patterns for a chapter of an expository text: either high-cohesion at the beginning (to benefit low-knowledge readers), followed by a gradual decline to lower cohesion at the end (to benefit the readers' growing knowledge); or, high-cohesion throughout the text.

In terms of difficulty, we predicted that readability levels may rise across the chapters of the expository texts analyzed. We predicted this to be particularly likely in the science genre, because science texts tend to increase in terms of complexity of ideas as they progress, and expressing complexity often requires longer sentences and longer words (Zipf, 1949).

If we add this idea of difficulty organization to our optimal cohesion distribution hypothesis, the most favorable results from our analyses of expository text would imply a steady increase in difficulty across chapters. The reverse, a decrease in difficulty, would represent the least optimal organization for students. That is, beginning a textual unit with high difficulty and low cohesion, suggesting that both writing style and topic material would be difficult to comprehend, would represent a distribution that could burden low-knowledge students.

We also predicted that results from the science domain would differ slightly from those of the history and narrative domains. As previously mentioned, science texts tend to cover less familiar subjects, meaning that greater redundancy is required to explain complex issues. We predicted that this need for redundancy would result in the science texts having higher cohesion levels than the history and narrative texts. Because sentence length and word length are not affected by redundancy in the same way, we did not predict significant differences between the science and history domains in terms of FKGL.

For narratives, however, we predicted that these non-expository texts would show lower levels in both cohesion and difficulty than science and history domains. Expository texts often require redundancy, resulting in higher cohesion, but they normally illustrate complex topics,

resulting in higher difficulty. Narratives differ from expository texts in the way that they normally involve more familiar topics and require less redundancy, resulting in a text with both lower difficulty and lower cohesion.

Method

In this study, we applied Coh-Metrix to seven commonly used high-school textbooks (three science, four history) and six frequently assigned classic novels to assess their textual organization in terms of cohesion and difficulty (see Appendix). The textbooks were sampled from a large corpus of school texts provided by MetaMetrics Inc. We obtained and viewed hard copies of the textbooks to ensure the paragraph and chapter breaks were correctly located. The hard copies also aided in removing picture captions, footnotes, and commentaries which were occasionally present in the electronic versions. We divided the texts into chapters, totaling 130 chapters for the science texts and 117 chapters for the history texts, and sampled three representative sections from each chapter of each textbook. We took the first 1000 words to represent the beginning, the last 1000 words to represent the end, and the 500 words on either side of the middle word of the chapter to represent the middle.

After viewing several high school reading lists from city and county school systems, we selected six novels which appeared at high frequencies. These six were also chosen based on number of chapters and chapter lengths. From these narratives we sampled 400 word sections from the beginning, the middle, and the end of each chapter of each book as representative sections of those parts of the chapters. These sections were shorter due to restrictions presented by the shorter length of the chapters.

While Coh-Metrix offers over 400 indices of cohesion and difficulty, many of these indices cover overlapping aspects. In certain studies, multiple comparable approaches are relevant, but for this study we required just four indices, representing various aspects of cohesion and difficulty, to suitably test our hypotheses. The four indices we used are described below, for a complete explanation of all Coh-Metrix indices, please see Graesser et al. (2004).

For cohesion, three indices were selected: argument overlap, latent semantic analysis (Landauer and Dumais, 1997), and connectives frequency. Argument overlap is a coreference index that determines the proportion of adjacent sentences that share one or more arguments (i.e., noun, noun phrase, pronoun). Latent semantic analysis differs from argument overlap, as values are produced for relative similarities as opposed to specific word overlap. For example, *desk/desks* scores more highly than *chair/table*, which in turn scores more highly than *pencil/alligator*. Connectives frequency, which identifies words like *and*, *or*, and *because*, serves to offer explicit reference cues to the reader. To represent a traditional measure of textual analysis, we employed Flesch-Kincaid

Grade Level. This index is determined by the text's average sentence length and the average number of syllables per word. The rationale of this index is based on research suggesting that longer sentences and longer words are more difficult to process (Graesser et al., 2001; Just & Carpenter, 1980; Zipf, 1949).

Results

Values for each of the variables chosen for analysis were calculated for each section of each chapter for the three science texts, four history texts, and six narrative texts on Flesch-Kincaid Grade Level, LSA similarity, argument overlap, and incidence of connectives. ANOVA's were performed to determine whether there were significant trends across chapters. The scores for the three sections of each chapter for the three genres are shown in Table 1.

Table 1. Measures of Cohesion and Difficulty in terms of Flesch-Kincaid Grade Level (F-K), Latent Semantic Analysis (LSA), Argument Overlap (AO), and Incidence of Connectives (Con) for the Beginning, Middle, and End of Chapters for Science, History, and Narrative Texts

Science

| | Beginning | Middle | End | Sig |
|-----|--------------|--------------|--------------|-----|
| F-K | 10.39 (0.12) | 10.63 (0.10) | 10.76 (0.12) | ** |
| LSA | 0.38 (0.01) | 0.39 (0.01) | 0.34 (0.01) | ** |
| AO | 0.70 (0.01) | 0.71 (0.01) | 0.67 (0.01) | ** |
| Con | 68.25 (1.03) | 65.45 (1.11) | 67.51 (0.91) | |

History

| | Beginning | Middle | End | Sig |
|-----|--------------|--------------|--------------|-----|
| F-K | 10.43 (0.07) | 10.57 (0.06) | 10.82 (0.07) | ** |
| LSA | 0.24 (0.01) | 0.23 (0.01) | 0.24 (0.01) | |
| AO | 0.43 (0.01) | 0.46 (0.01) | 0.45 (0.01) | * |
| Con | 68.88 (0.66) | 69.40 (0.69) | 69.03 (0.60) | |

Narrative

| | Beginning | Middle | End | Sig |
|-----|--------------|--------------|--------------|-----|
| F-K | 9.38 (0.17) | 8.06 (0.18) | 7.95 (0.18) | ** |
| LSA | 0.31 (0.01) | 0.30 (0.01) | 0.31 (0.01) | |
| AO | 0.42 (0.03) | 0.39 (0.03) | 0.39 (0.03) | ** |
| Con | 93.53 (6.17) | 91.80 (6.05) | 92.94 (6.13) | |

Note: Standard errors appear in parentheses,

* $p < .05$; ** $p < .01$

Flesch-Kincaid Grade Level The science textbooks showed an increase in grade level across each chapter, $F(2,124) = 3.36, p < 0.01$. Planned comparison contrasts showed that this increased monotonically from the beginning to the end of each chapter. The history textbooks showed the same pattern, a linear increase across the three sections of each chapter, $F(2,422) = 13.44, p < 0.01$. The narrative texts, however, displayed a linear decrease in grade level difficulty across chapters, $F(2,448) = 13.62, p < 0.01$. There was no significant difference between the FKGL values for history and science texts, but the FKGL values for narrative texts were significantly lower than both science and history domains. These results were in

line with our predictions that difficulty would increase across the course of an expository chapter. These results were also in line with our expectation that there would be no significant difference in difficulty between the expository domains.

LSA The LSA score for each chapter section was calculated by calculating the average cosine between each sentence in the section and the section as a whole. Higher LSA cosines imply higher cohesion. For the science texts, the results partially mirrored the findings for FKGL showing a slight decrease in cohesion (suggesting an increase in difficulty for this genre) across chapters, $F(2,124) = 10.23, p < 0.01$. However, unlike the effect for FKGL, there is no meaningful distance between the beginning and middle of the chapters: planned comparison contrasts revealed that the only significant difference appeared between the first two sections and the final section. The science texts received the highest LSA values, the history texts averaged lowest, and the narrative texts fell in between. The higher cohesion levels for the science texts concurred with our initial predictions.

Argument Overlap For the science texts, argument overlap performed similarly to LSA, showing an overall downward trend in cohesion across chapters, $F(2,124) = 5.11, p < 0.01$. Further emulating the LSA effect, argument overlap found no difference between the science texts' beginning and middle sections, but a significant difference between the first two sections and the final section. For the history texts, however, argument overlap found an effect in the opposite direction: an increase in argument overlap across the course of a chapter, $F(2,422) = 6.25, p < 0.01$. Contrasts showed that the only difference was between the beginning and middle sections. For the narrative texts, argument overlap showed a decrease in cohesion across chapters, $F(2,448) = 5.01, p < 0.01$. Planned comparisons indicated that there were significant differences comparing the beginning to the middle and end. As can be seen from Table 1, history and narrative chapters had lower argument overlap than science chapters, reflecting the findings of genre differences with LSA. This was in line with our prediction that narrative and history texts would have similar cohesion levels.

Connectives There were no differences in the use of connectives across sections of a chapter within any of the three domains.

Discussion

The trends we observed for the four indices within and between the three genres were in line with our predictions. The two expository domains, science and history, did not differ significantly in their levels of Flesch-Kincaid difficulty. However, the LSA and argument overlap measures indicated that the science texts were higher in cohesion than the history texts. The findings were in line

with our predictions, which were based on a simple comparison of the general characteristics and purposes associated with each genre. Specifically, history texts tend to follow simple narrative patterns with chronological development and relatively familiar themes; science texts, on the other hand, normally explore less familiar topics.

While the science and history texts both displayed an increase in difficulty across a chapter, the narrative texts steadily decreased in difficulty from the beginning to the end of a chapter. This result, in addition to the finding that narrative texts were assigned lower overall grade levels than expository texts lends support to the claim that narrative and expository texts are structured quite differently.

In terms of optimal design for distributions of cohesion and difficulty, the results showed that both cohesion and difficulty approaches have utility in evaluating texts for appropriateness. FKGL scores showed a linear increase in difficulty across expository chapters and a linear decrease in difficulty in narrative chapters. These results suggest that texts for both expository domains gradually rise in complexity as they develop. It also suggests that chapters in narrative fiction have structural characteristics that differ from textbooks. Topic material in science texts seems to be introduced slowly, with simpler, more readable writing early on in a chapter.

For the narratives, we speculate that FKGL values may indicate, not the difficulty of the text, but the speed of pacing. Material in a novel seems to be introduced more slowly at the beginning, the author taking time to establish character and setting. Because there can often be less dialogue in these beginning sections, the sentence lengths tend to be longer, causing the FKGL score to rise. The descriptions and setting of scene that typically occurs at the beginning of narrative chapters often calls for low frequency, multi-syllabic words, whereas advancing the storyline, usually taking place in the middles and ends of chapters, can involve less complex language.

While the cohesion scores for the science texts did not show a linear decrease, we must remember that we cannot assume that students constantly gain knowledge as they read through a text. Thus, we proposed that an optimal design for expository texts is that they either gradually fall in cohesion or that they stay more or less constantly high. The results of this study suggest that cohesion levels for science texts stay fairly steady across the text with a slight tailing off of cohesion towards the end. For the history texts, however, there is some evidence that cohesion rates actually rose toward the end of the chapters. According to our hypothesis of optimal cohesion distribution in expository texts, this organization does not adhere to either of the two proposed optimal patterns.

Our future research will continue to pursue these notions. We intend to gather a greater number of textbooks and examine differences in cohesion and difficulty as a function of the assigned grade level of the textbook. We

also intend to gather a larger and more diverse sample of narratives. Our overarching goal is to further our understanding of structural organization in text for a variety of genres, particularly the science, history, and narrative domains.

Acknowledgements

This research was supported by the Institute for Education Sciences (IES R3056020018-02). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the IES.

References

- Baayen, R.H., Piepenbrock, R., & van Rijn, H. (Eds.) (1993). The CELEX Lexical Database [CD-ROM]. University of Pennsylvania, Philadelphia: Linguistic Data Consortium.
- Beck, I.L., McKeown, M.G., Sinatra, G.M., & Loxterman, J.A. (1991). Revising social studies text from a text processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly*, 26, 251-276.
- Best, R., Floyd, G., & McNamara, D. (2005). Differential aptitudes contributing to children's comprehension of narrative and expository texts. *Reading Psychology*. Submitted for publication.
- Bowen, B.A. (1999). Four puzzles in adult literacy: Reflections on the national adult literacy survey. *Journal of Adolescent and Adult Literacy*, 42, 314-323.
- Britton, B.K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83, 329-404.
- Coltheart, M. (1981). The MRC psycholinguistic database quarterly. *Journal of Experimental Psychology*, 33A, 497-505.
- Graesser, A.C. (1993). Inference generation during text comprehension. *Discourse Processes*, 16, 1-2.
- Graesser, A.C., Karnavat, A.B., Daniel, F.K., Cooper, E., Whitten, S.N., & Louwerse, M. (2001). A computer tool to improve questionnaire design. In *Statistical Policy Working Paper 33, Federal Committee of Statistical Methodology* (pp. 36-48). Washington, DC: Bureau of Labor Statistics.
- Graesser, A.C., McNamara, D., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers*, 36, 193-202.
- Jurafsky, D., & Martin, J. (2000). *Speech and language processing*. Englewood, NJ: Prentice Hall.

- Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Klare, G.R. (1974-1975). Assessing readability. *Reading Research Quarterly*, 10, 62-102.
- Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.
- Linderholm, T., Everson, M.G., van den Broek, Mischinski, M., Crittenden, A., & Samuels, J. (2000). Effects of causal text revisions on more and less skilled readers' comprehension of easy and difficult text. *Cognition and Instruction*, 18, 525-556.
- Louwerse, M.M., McCarthy, P.M., McNamara, D.S., & Graesser, A.C. (2004). Variation in language and cohesion across written and spoken registers. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 843-848). Mahwah, NJ: Erlbaum.
- McCarthy, P.M., Graesser, A.C., & McNamara, D.S. (2006, July). Distinguishing genre using Coh-Metrix indices of cohesion. *Paper presented at the Society for Text and Discourse conference*, Minneapolis, MN.
- McCarthy, P.M., Briner, S.W., Rus, V., & McNamara, D.S. (in press). Textual Signatures: Identifying text-types using Latent Semantic Analysis to measure the cohesion of text structures. In: A. Kao, & S. Poteet (Eds.). *Natural Language Processing and Text Mining*. London: Springer-Verlag UK.
- McNamara, D.S., & Kintsch, W. (1996). Learning from Text: Effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247-287.
- McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1-43.
- McNamara, D.S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, 55, 51-62.
- McNamara, D.S., Louwerse, M.M., & Graesser, A.C. (2006). Coh-Metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension. *Annual Report to the Institute of Educational Sciences*, R305G020018.
- Miller, G.A., Beckwith, R., Feldbaum, C., Gross, D., & Miller, K. (1990). *Five papers on WordNet* (Tech. Rep.). Princeton University, Cognitive Science Laboratory.
- O'Reilly, T., & McNamara, D.S. (in press). Good texts can be better for skilled comprehenders. *Discourse Processes*.
- Ozuru, Y., Dempsey, K., Sayroo, J., & McNamara, D.S. (2005). *Effects of text cohesion on comprehension of biology texts*. Proceedings of the 27th Annual Meeting of the Cognitive Science Society (pp. 1696-1701). Hillsdale, NJ: Erlbaum.
- Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.
- Tonjes, M.J., Ray, W., & Zintz, M.V. (1999). *Integrated Content Literacy*. New York: The McGraw-Hill Publishers.
- Zipf, G.K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.

Appendix

- Appleby, J., Brinkley, A., Broussard, A.S., McPherson, J.M., & Ritchie, D.A. (2003). *The American Republic Since 1877*. New York, NY: Glencoe/McGraw-Hill.
- Austin, J. (1813). *Pride and Prejudice*. New York, NY: Washington Square Press.
- Boyer, P. (2001). *The American Nation*. Austin, TX: Holt, Rinehart, and Winston.
- Brönte, C. (1847). *Jane Eyre*. Cornhill: Smity Elder and Company.
- Cather, W. (1918). *My Ántonia*. Boston, MA: Houghton Mifflin.
- Danzer, G.A., Klor de Alva, J.J., Woloch, N., & Wilson, L.E. (2002). *The Americans: Reconstruction through the 20th Century*. Evanston, IL: McDougal Littell.
- Davis, M. (1999). *Modern Chemistry*. Austin, TX: Holt, Rinehart, & Winston.
- Dickens, C. (1859). *A Tale of Two Cities*. London: Chapman Hall.
- Hawthorn, N. (1850). *The Scarlet Letter*. Boston: Ticknor, Reed & Fields.
- Myers, R.T., Oldham, K.B., & Tocci, S. (2000). *Chemistry: Visualizing Matter*. Austin, TX: Harcourt Brace and Company.
- Nash, G.B. (1999). *American Odyssey*. Westerville, OH: Glencoe/McGraw-Hill.
- Serway, R.A., & Faughn, J.S. (1999). *Holt Physics*. Austin, TX: Holt, Rinehart, and Winston.
- Twain, M. (1884). *The Adventures of Huckleberry Finn*. New York, NY: Washington Square Press.