ARTICLES

Readability, Logodiversity, and the Effectiveness of College Science Textbooks

Rebecca S. Burton

Department of Biology, Alverno College, PO Box 343922, Milwaukee, WI 53234-3922

Rebecca.burton@alverno.edu

Abstract: Textbooks are required in most introductory college science courses, but students may not be benefitting from the textbooks as much as their instructors might hope. Word use in the textbooks may influence textbook effectiveness. I tested whether either the amount of technical vocabulary or the readability had a significant effect on students' ability to learn general biology concepts. I provided different versions of the same reading, then tested students on the content. On the topic with the lowest overall post-reading quiz scores, students who received readings with less technical vocabulary outperformed their peers (P = 0.03). Textbooks did not appear to be an important source of learning for students in this study; fewer than half the students reported that they were reading the assigned chapters near the start of the semester, and this number declined sharply. Students had difficulty correctly answering questions immediately after reading brief selections, indicating a low level of comprehension. Changes in textbooks and teaching strategies may improve student learning and reading compliance.

Key words: readability, logodiversity, textbooks

INTRODUCTION

Unless students can learn from their textbooks, there is little purpose in requiring them. Overwhelmingly, introductory college science courses use textbooks, but questions remain about what aspects of the actual prose result in greater student learning of scientific concepts. When students enter college from high school, they encounter textbooks that are more difficult to read and may find that their literacy skills are inadequate (Williamson, 2008). There is no guarantee that students are actually reading at the level that corresponds to their years of education. A national study (Baer et al., 2006) revealed that only 38% of students nearing graduation from 4-year colleges were proficient enough in prose literacy to understand materials such as textbooks, and their literacy improved very little during their time in college.

Increasing the diversity of people entering scientific careers is widely recognized as an important goal, yet literacy is likely to be lower for college students from some ethnic groups and those for whom English is not the primary language (Baer et al., 2006). Students with learning differences related to reading will likely have additional problems with comprehension. It is reasonable to assume that student success in biology may depend—at least in part—on the ability to read textbooks effectively, and that many biology students lack sufficient ability. Instructors may wish to select textbooks that are effective with students who are poor readers, but what makes a textbook readable? A number of indices have been developed to measure the readability of written materials in terms of their syntactic and semantic difficulty (Fry, 2002). Most of these grade level scores are based on length of sentences and length of words. These indices are generally intended for use in evaluating materials for K-12 schools or for the general public. Making sentences shorter and simpler will result in a lower grade level score, but may not make college biology textbooks easier for high school seniors to understand (Johnson and Otto, 1982). This may be because the shorter sentences can make it harder for students to see the logical relationships between adjacent sentences (Armbruster et al., 1985). In a study of college business students, the use of more readable textbooks was correlated with greater student retention and higher grades (Spinks and Wells, 1993). Landrum et al. (2012) found a similar pattern with psychology students.

In addition to word and sentence length, unfamiliar vocabulary can be a challenge for readers. College textbooks can differ greatly in logodiversity, a measure of how many technical terms are introduced and how often they are used (Burton, 2011). Science contains so many technical terms that scientists in different fields often have difficulty communicating. If students are to learn science, they must have technical language translated for them (Montgomery, 2004). Some mastery of technical vocabulary is essential for becoming biologically literate, but there is no widespread agreement on which words are critical at the introductory level. Some researchers have expressed concern about the number of technical terms and volume of information included in recent textbooks (Blystone, 1987; Lord, 2007). A balance should be found so that students are introduced to enough vocabulary to allow them to

Table 1. Calculations for three measures of readability. For each, the index is intended to correspond with the number of years of schooling required to comprehend the writing.

The Flesch–Kincaid Grade Level (FKGL) was calculated as:

$$0.39\left(\frac{total\ words}{total\ sentences}\right) + 11.8\left(\frac{total\ syllables}{total\ words}\right) - 15.59$$

The Coleman-Liau Index (CLI) was calculated as ***0.0588L – 0.296S – 15.8***, where L is the average number of letters per 100 words and S is the average number of sentences per 100 words.

The Automated Readability Index (ARI) was calculated as:

$$4.71\left(\frac{characters}{word}\right) + 0.5\left(\frac{words}{sentence}\right) - 21.43$$

communicate effectively, but not so much that they are overwhelmed.

The purpose of this experiment was to examine factors that might influence the effectiveness of college biology textbooks. The study compared comprehension of students who read textbook excerpts that differed from the original in one of two ways: logodiversity or grade level index. It also examined students' compliance with assigned readings and whether reading the textbook before class improved their performance.

MATERIALS AND METHODS

Students' comprehension of readings was analyzed by presenting students in an introductory majors biology course with one of three versions of a reading in biology. One version consisted of the original wording from a ubiquitous textbook. Another version was altered to reduce the grade level score by using shorter words and sentences. The third was altered to reduce logodiversity by replacing technical terms with less specialized vocabulary. Students then completed a brief quiz based on the readings. This was repeated for a total of four topics spread over a semester.

Readings

On four occasions, students received a brief (approximately one-page) reading on the topic to be covered in class that day. They then answered a series of questions. The topics were presented in this order: population dynamics, basic enzyme properties, the light-dependent phase of photosynthesis, and Mendelian inheritance. These topics were chosen mostly to allow the sessions to be spaced at roughly equal intervals (approximately session per month) throughout the course. All four original readings were selected from Campbell's *Biology: Concepts & Connections* (Campbell et al., 2006). This textbook was chosen because of its popularity and because it uses a large number of technical terms, has a relatively complex sentence structure, and uses a rich non-technical vocabulary. Sections were edited to remove references to illustrations and arranged so they covered the topic of interest, but sentences were left intact for the readings termed *original*.

Two other versions were created from each original reading. In one version, sections were edited to remove some of the technical vocabulary. These were termed *low logodiversity*. For example, "F1" was replaced by "first generation." Sentence structure was changed only to preserve meaning when the number of technical words was reduced. The decisions regarding word substitution were subjective.

In another version, sentence structure and words were changed to make the reading difficulty of the section more appropriate for students with lower reading ability. These were termed *low grade level*. Technical vocabulary was usually left intact, though a shorter technical term was sometimes substituted for a longer one. The decisions regarding sentence structure were subjective, but the effect of the changes was evaluated with readability indices. The readability indices used were the Flesch–Kincaid Grade Level (FKGL), the Coleman-Liau Index (CLI) and the Automated Readability Index (ARI). All of these indices provide a number that is meant to correspond to the level of education needed to comprehend the reading (Table 1). In these measures, a first-year college student generally is assumed to be able to comprehend readings with scores of about13 or lower.

Because the weighting of sentence length versus word length in determining the readability level varies from index to index, the three indices tend to give slightly different scores for identical readings, but the rank of the scores was consistent for each set. The original version had the highest score (meaning it would be the most challenging to read; Table 2), the version with reduced logodiversity tended to score approximately one grade level lower, and the version edited to reduce the grade level readability scores tended to be an additional grade lower (meaning it should be understandable to a person with two fewer years of education than those required to read the original).

Subjects

The students in this study were all enrolled in three of the six sections of the introductory biology course for science and nursing majors at Alverno College in the spring semester of 2011 (approved by

Table 2. Readability indices for the three versions of each of the readings. The index is assumed to correspond to the level of education needed to understand the passage. FKGL = The Flesch-Kincaid Grade Level, CLI = Coleman-Liau Index, ARI = Automated Readability Index.

Subject	Version	FKGL	ARI	CLI	Average
Population Dynamics	Low Grade Level	10.75	11.05	13.83	11.88
	Low Logodiversity	12.00	12.96	14.42	13.13
	Original	13.38	14.09	15.64	14.37
Enzymes	Low Grade Level	8.05	8.60	11.49	9.38
	Low Logodiversity	9.44	10.37	11.79	10.53
	Original	10.27	11.26	12.57	11.37
Photosynthesis	Low Grade Level	9.42	9.28	12.38	10.36
	Low Logodiversity	10.53	10.70	12.44	11.23
	Original	11.58	11.82	13.51	12.30
Inheritance	Low Grade Level	7.18	7.27	9.89	8.11
	Low Logodiversity	9.51	10.47	11.18	10.39
	Original	10.78	11.61	11.92	11.43

the Alverno Institutional Review Board, IRB-011M-10). Two sections were taught by the author, the other by another faculty member in the biology department. There is no reason to believe that the students in the three sections used in this study differed in important ways from those in the other sections, but because sections were not compared with each other, this factor would not be expected to influence the outcomes of the experiment. Students typically take this course in the second semester of their first year, and must complete a course in introductory physical science as a prerequisite. All Alverno College undergraduates are female. In spring of 2011, 58% of the total undergraduate student body identified as Caucasian American, 17% identified as African American, 14 % identified as Hispanic American, 5% identified as Asian American, and 4% identified as non-US residents or reported multiple ethnicities. Of the full-time undergraduate students, 88% were awarded some type of financial aid (Source: Alverno Institute). In order to preserve anonymity, demographic information was not collected from students in this study.

Testing

Testing occurred immediately before the specific topics were covered in class. Students did not know in advance what topics would be part of the study. At the beginning of class, the instructor read a brief script about the study, informing students that they could opt out of it with no penalty and all responses (including their willingness to participate) would be anonymous. Each student was given one reading (approximately one page long), one question sheet, and one answer form. The three versions of the reading were distributed randomly and had been marked A, B, or C to identify the version. The letters were used randomly each time. For example, the original was labeled C one time and A another. All students received the same quiz. Students were provided with time in class (approximately 20 minutes) to complete the reading and answer the questions. Students were instructed to answer the questions without looking at their textbook or the reading they had just completed. Instructors observed no attempts at non-compliance with this. Testing was done for all sections at the same point in the course—the first day the topic would be covered—but not the same calendar day because the sections met on different days of the week. There was no evidence of increased reading compliance or increased scores to suggest that students had been alerted to prepare for the testing by their peers.

Students answered the questions using Immediate Feedback Assessment Technique® forms (IF-AT® forms, Epstein Educational Enterprises, Cincinnati, Ohio). These forms use scratch-off squares similar to those used on many lottery tickets. A student scratches off the square corresponding to the response she believes is correct. If it is correct, she will see a star revealed in the square; if not, she can review the question and possible answers to make another selection until she has found the correct answer. The student, therefore, had instant feedback on her performance. She could mark the correct answers on her question sheet and use this information for future study. This method also informed the researcher how many attempts the student needed to get the correct answer when given four choices. Students had used IF-AT® forms previously in review activities.

All students completed the reading and quiz as a learning experience. The first question asked the student whether she were willing to have her anonymous answer form used in the study. Only one student on one quiz requested that her answers not be used in the study.

The second question asked which reading selection she had received that day (A, B, or C). The third asked whether she had read the assigned relevant chapter in her own textbook (*Biology: A Guide to the Natural World* by Krogh) before coming to class. For the first three questions, students were

told to ignore whether their scratch-off revealed a star.

The remaining six questions tested comprehension of the subject; five were based on the in-class reading. One question tested vocabulary recall, one tested factual recall, another required the student to choose a correct summary of the concept, another required the student to apply the concept, and an extension question required the student to reason beyond the facts provided (e.g., to predict what would happen to the proportion of recessive genes over many generations). Naturally, there was overlap in the questions. For example, a vocabulary question was also a test of simple recall.

A control question asked about information that was in the assigned chapter on the topic, but that could not be extrapolated from the information in the in-class reading. The order of the question types was scrambled so that, for example, the simple recall question was a different number on each of the four quizzes in the study.

Analysis

The IF-AT® forms allow multiple possible points for each question. A question that was answered correctly on the first try (only one box scratched off) was scored as 5 points. A correct answer on the second try was scored as 3 points. A correct answer on the third try received 1 point. If all boxes were scratched off, no points were awarded.

The word "total" is used below to reflect the point total on the five questions that could be answered correctly using the information in the one-page reading provided in class. The score from the control question was used in other analyses, but not included in the total.

ANOVA tests were used to determine whether student performances on the quizzes were influenced by the type of reading they had done (original, low technical vocabulary, or low grade level index). One-tailed independent T-tests were used to compare students who reported completing the assigned reading with those who reported either that they had not or were not sure. All statistical analyses were performed using Microsoft® Excel® 2007.

RESULTS

In general, a reduction in logodiversity resulted in significantly improved scores for one topic, but not others. Compliance with pre-class reading assignments was low, and did not measurably influence student performance on the quizzes.

The three versions of the population ecology reading, the first of the four, did not result in any differences in student scores (Fig. 1, P = 0.8). Performance was strong overall (average = 23.32 of 25 possible, with a low score of 15). About half the students reported not having completed the assigned reading before class (26 reporting having read it, 27 reported not reading it, and 3 were unsure). There

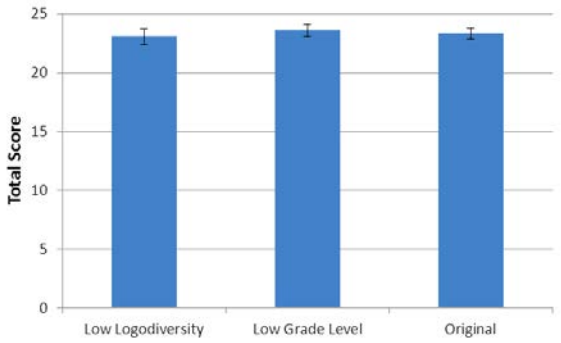

Fig. 1. Mean (+/- standard error) scores for students on a population ecology quiz with a maximum possible score of 25. Quizzes were given immediately after reading selections that differed in terms of grade level reading index or logodiversity index. Sample size = 19, 19, and 18. ANOVA, P = 0.8.

was no significant difference in performance between those who had done the reading before class and those who had not (unpaired T-test, 1-tailed, P = 0.089).

By contrast, the type of reading selection had a significant effect on student performance on the enzyme quiz (P = 0.03, Fig. 2). Those who read the version with a low logodiversity performed better on the quiz than those who read the other selections. Those who read the original tended to fare the worst. The overall performance on this quiz was the lowest of the four (average = 16.23, with a low score of 7). Students who had completed the assigned reading did no better than those who had not (unpaired T-test, 1-tailed, P = 0.28, N = 52).

Student performance on the photosynthesis quiz was not influenced by the reading selection type (Fig. 3, P = 0.88). Performance was moderate overall (average = 18.81 of 25 possible, with a low score of 9). Of the 52 students taking this quiz, only 15 reported having read the assigned chapter before attending class. These students did not perform significantly better than their peers (unpaired T-test, 1-tailed, P = 0.404).

The results of the quiz on Mendelian genetics were a bit different from the other quizzes in that

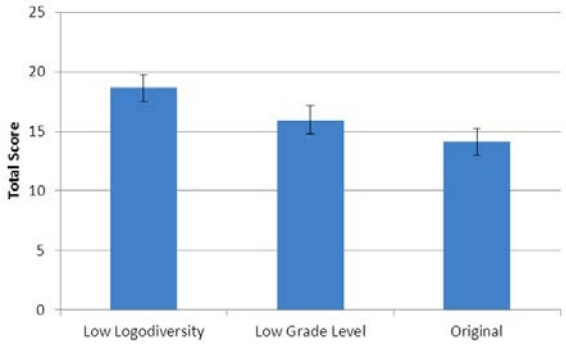

Fig. 2. Mean (+/- standard error) scores for students on an enzyme quiz with a maximum possible score of 25. Quizzes were given immediately after reading selections that differed in terms of grade level reading index or logodiversity index. Sample size = 17, 17, and 18. ANOVA, P = 0.03.

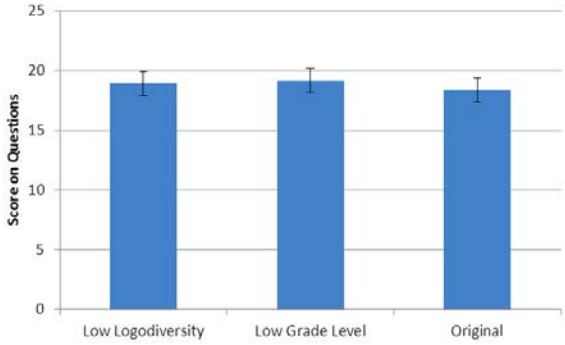

Fig. 3. Mean (+/- standard error) scores for students on a photosynthesis quiz with a maximum possible score of 25. Quizzes were given immediately after reading selections that differed in terms of grade level reading index or logodiversity index. Sample size = 14, 14, and 15. ANOVA, P = 0.08.

students who read the low logodiversity selection tended to perform less well than their peers, but the difference was not significant (ANOVA, P = 0.067, Fig. 4). Overall, performance on this quiz was slightly lower than that of the photosynthesis quiz, though still higher than performance on the enzyme quiz (average = 17.7, low score = 8, N = 48). The number of students completing the pre-class reading assignment was so low that a comparison between the students who had done the reading and those who had not (5 vs. 42) would not be valid.

Compliance with course reading assignments was low at the beginning of the semester and declined sharply throughout the semester (Fig. 5). Less than half of the students were completing the readings near the start of the semester, and scarcely more than 10% were reading the assigned chapter by the tenth week of the semester. However, completing the assigned reading did not lead to better performance on any of the questions, including the control questions.

Overall, students performed best on the question in each quiz that asked them to summarize the reading selection, but there was variation (Table 3). They also generally performed well on simple recall. As would be expected, they tended to perform poorly on the control questions.

DISCUSSION

Many college students may find themselves unprepared to read their college science textbooks effectively (Baer et al., 2006; Williamson, 2008). The density of the information is probably one source of difficulty (Smith et al., 2010), but the writing itself may make the task more difficult by using technical vocabulary, complex wording, or words that are not part of the students' working vocabulary. In this study, reducing the technical vocabulary of passages from a college biology textbook increased student reading comprehension of the topic to a limited degree. For the topic with which students had the most difficulty (enzymes), reduction in technical vocabulary, general vocabulary, and sentence length may have allowed students to focus on more difficult

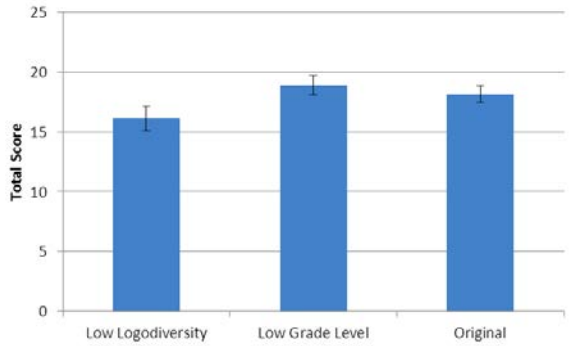

Fig. 4. Mean (+/- standard error) scores for students on a transmission genetics quiz with a maximum possible score of 25. Quizzes were given immediately after reading selections that differed in terms of grade level reading index or logodiversity index. Sample size = 14, 16, and 17. ANOVA, P = 0.067.

concepts. For the other topics, there was no significant difference in the performance of the students who read different versions of the passages. If students found this concept fairly easy to understand, they may have been able to tolerate writing with higher logodiversity and longer words and sentences.

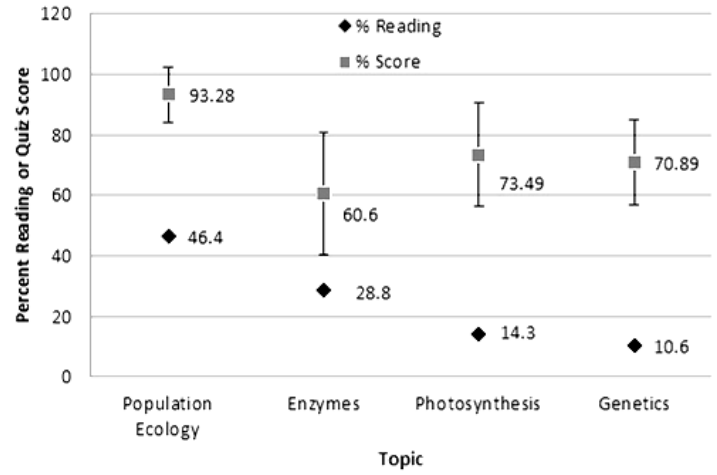

Fig. 5. Percent of students who reported having completed the assigned reading before coming to class and average percent score (+/- standard deviation) on the quiz for each topic for all students. Topics are arranged in chronological order.

Table 3. Scores on different types of questions.

Topic	Recall	Vocabulary	Summary	Extension	Application	Control
Population	4.89	4.82	4.68	4.45	4.48	3.41
Enzymes	2.27	3.90	3.08	4.15	2.83	3.08
Photosynthesis	4.51	2.86	4.56	2.14	4.30	2.77
Genetics	4.00	2.55	4.32	4.11	2.74	2.94
Mean	3.92	3.53	4.16	3.71	3.59	3.05

Whether students can comprehend a reading may be a function of the interaction between content difficulty and the level of the writing itself. Wright and Spiegel (1984) found that high school teachers' predictions of readability were more accurate than the Fry readability index at predicting student comprehension of biology textbook readings. Apparently, teachers considered the difficulty of the subject matter, rating a reading on gene structure less readable than an introduction to ecology, even though the readings had equal scores on the Fry scale. In the current study, students performed very well on the ecology quiz and the version of the reading made no difference in their performance. They had more difficulty with genetics. There was a trend toward the students who received the low logodiversity reading performing worse than their peers, but it was not significant. The lack of Punnett squares and other diagrams in the in-class reading may have made the concepts so difficult to understand that the level of the writing was inconsequential. Illustrations can be critical in helping students learn biological concepts (Butcher and Kintsch, 2004; Rybarczyk, 2011).

There may be other reasons the reduction in logodiversity and reading level did not help the students on three of the topics. Perhaps the changes in the selections made some concepts more difficult to understand, but not in ways that the indices measured. Another possibility is that the reductions in reading level were not sufficient to make a difference. Students with significant challenges - such as those with poor academic preparation, learning differences, and those for whom English is not a primary language - may have benefitted from a greater reduction in the complexity of the readings.

Throughout this study, students who completed the reading before class were outnumbered by those who did not. By the end of the study almost no students were still doing the "required" reading. Other research has also found that students tend to complete very little of the assigned textbook reading (Bonner and Holliday, 2006; Brost and Bradley, 2006). Students may be using their textbooks as review tools, sources of diagrams, or as resources for discussions and homework, but not reading entire chapters to prepare for class. This attitude may persist into early college years. It is worth asking whether students would be more likely to read textbooks that had simpler prose. Landrum et al. (2012) found that college psychology students preferred textbooks that were more readable, that students were more likely to actually read these books, and that those doing the reading performed better. Brost and Bradley (2006) found that the difficulty of a reading may be secondary to faculty behavior; if students believe that reading before class will not help them--because readings will be summarized or ignored in class--reading compliance is likely to be low.

When readers are new to a discipline, they lack not only content knowledge, but also cognitive frameworks for identifying key concepts and making sense of the information presented. They may also lack intrinsic motivation to expend the effort needed to learn from complex text (Jetton and Alexander, 2000). Faculty may not be considering these factors in their textbook selection. As Crow (2004) and Dutch (2005) point out, the more engaging books are not the books that instructors tend to choose for majors-only courses. Selecting textbooks with appropriate complexity of prose and other factors that influence student learning might improve reading compliance and effectiveness.

In this study, doing the pre-class reading did not appear to have been particularly helpful to students. One could argue that the influence of the in-class reading was so strong that it swamped any differences between students who had completed the pre-class reading and those who had not, but this lack of pre-class reading advantage was also found on the control question—the one that could not be answered from information in the in-class reading. While this was only one question per topic, the pattern was consistent through all four topics. Perhaps students in this study were reading the chapters superficially without really comprehending the concepts. Lord (2007) suggests that reading before lecture can result in more misconceptions, or in the ability to repeat information without understanding it. This could lead to students overestimating their understanding of what they read. Norris, Phillips and Korpan (2003) found that university students tended to rate media reports of scientific discoveries as easy to read, even when they failed to understand key features of the articles. Students in their study seemed to believe that if they could understand the words, they understood the article. In actuality, the students were unable to perform higher order tasks that were critical to comprehension of the readings, but failed to recognize this.

Some research suggests that instructors may be able to increase their students' effectiveness at reading by using questions to probe understanding.

Smith et al. (2010) found that a questioning strategy improved student comprehension of a passage on physiology. Posing questions that require higher level thinking may increase student understanding (Lord, 2007; Pestel, 1997). Reading quizzes can hold students accountable for learning from the assigned text. Course management systems can provide quizzes without using class time or faculty grading effort. Students can be allowed to access questions before they read, which can assist them to focus on key ideas. Questions can rely on higher order processes because students can be given more time and scaffolding than in a brief quiz at the start of class. Many quiz programs provide answer-specific feedback options, allowing the instructor to explain why the answer is incorrect, give hints about avoiding common misconceptions, or explain in more detail the implications of a correct answer.

Asking questions may be as useful as answering them. Henderson and Rosenthal (2006) suggest having students ask the instructor questions between reading and attending class. Instructors may not be able to imagine how students think about a topic, and this method would provide insight on student understanding and misconceptions. Teaching students how to ask good questions involving analysis, evaluation, and prediction may help students use the textbook effectively. Marbach-Ad and Sokolove (2000) found that students in active learning courses can be taught to ask higher level questions by being introduced to criteria for evaluating questions. They suggest having students prepare questions in advance to avoid low quality, last minute questions. Requiring online submissions by a deadline would encourage advance preparation of questions.

Understanding our students' experience in reading textbooks can improve instruction. Attention to the reading level of materials, teaching effective reading strategies, and holding students accountable for completing readings can improve student compliance with reading assignments and may improve their comprehension of what they read.

REFERENCES

ARMBRUSTER, B.B., OSBORN, J.H. AND DAVISON, A.L. 1985. Readability formulas may be dangerous to your textbook. *Educational Leadership*, 42(7): 18-20.

BAER, J.D., COOK, A.L. AND BALDI, S. 2006. The Literacy of America's College Students. Washington, DC: American Institutes for Research.

BLYSTONE, R.V. 1987. College introductory biology textbooks: an important communicative tool. *American Biology Teacher* 49(7): 418-25.

BONNER, J.E. AND HOLLIDAY, W.G. 2006. How college students engage in note-taking strategies. *Journal of Research in Science Teaching* 43: 786-818.

BROST, B.D., AND BRADLEY, K.A. 2006. Student compliance with assigned reading: A case study: *Journal of Scholarship of Teaching and Learning* 6(2): 101-11.

BURTON, R.S. 2011. Bridges or barriers: Analysis of logodiversity in college biology textbooks. *Bioscene* 37(1): 3-7.

BUTCHER, K.R. AND KINTSCH, W. 2004. Learning with diagrams: Effects on inferences and the integration of information. *Journal of Educational Psychology* 98(1): 182-97.

CROW, L. 2004. The good, the bad, and the ugly: Introductory biology textbooks. *Journal of College Science Teaching* 33(4): 54-56.

DUTCH, S.I. 2005. Why textbooks are the way they are. *Academic Questions* 18(3): 34-48.

FRY, E. 2002. Readability versus leveling. *The Reading Teacher* 56(3): 286-91.

HENDERSON,C. AND ROSENTHAL, A. 2006. Reading questions: Encouraging students to read the text before coming to class. *Journal of College Science Teaching* 35(7): 46-50.

HOLLIDAY, W.G. AND CAIN, S.D. 2012. Teaching science reading comprehension: A realistic, research-based approach. Pp. 1424-1436 in Fraser, B.J., Tobin, K. and McRobbie, C.J. (eds.) Second International Handbook of Science Education. Springer: Netherlands.

JETTON, T.L., AND ALEXANDER, P.A. 2000. Learning from text: A multidimensional and developmental perspective. Pp. 285-310 in Kamil, M., Mosenthal, P., Pearson, P.D., and Barr, R. (eds.) Handbook of Reading Research, Volume III. Erlbaum, New Jersey.

JOHNSON, L.L., AND OTTO, W. 1982 . Effect of alterations in prose style on the readability of college texts. *Journal of Educational Research* 75(4): 222-29.

LORD, T. 2007. Please *don't* read the text before coming to lecture. *Journal of College Science Teaching* 37(1): 52-54.

LANDRUM, R.E., GURUNG, R.A.R, SPANN, N. 2012. Assessments of textbook usage and the relationship to student course performance. College Teaching 60: 17-24.

MARBACH-AD, G., AND SOKOLOVE, P.G. 2000. Can undergraduate biology students learn to ask higher-level questions? *Journal of Research in Science Teaching* 37(8): 854-70.

MONTGOMERY, S. 2004. Of towers, walls, and fields: perspectives on language in science. *Science* 303: 1333-35.

NORRIS, S.P., PHILLIPS, L.M. AND KORPAN, C.A. 2003. University students' interpretation of media reports of science and its relationship to background knowledge, interest, and reading difficulty. *Public Understanding of Science* 12: 123-45.

RYBARCZYK, B. 2011. Visual literacy in biology: A comparison of visual representations in textbooks and journal articles. *Journal of College Science Teaching* 41(1): 106-14.

SMITH, B.L, HOLLIDAY, W.G., AUSTIN, H.W. 2010. Students'comprehension of science textbooks using a question-based reading strategy. *Journal of Research in Science Teaching* 47(4) 363-379.

.

SPINKS, N. AND WELLS, B. 1993. Readability: A textbook selection criterion. *Journal of Education for Business* 69(2): 83-88.

WILLIAMSON, G.L. 2008. A text readability continuum for postsecondary readiness. *Journal of Advanced Academics* 19(4): 602-32.

WRIGHT, J.D. AND SPIEGEL, D.L. 1984. How important is textbook readability to biology teachers? *American Biology Teacher* 46(4): 221-25