When Do Students “Learn-to-Comprehend” Scientific Sources?: Evaluation of a Critical Skill in Undergraduates Progressing through a Science Major†

Tamara L. Marsh, Merrilee F. Guenther, and Stacey L. Raimondi*
Department of Biology, Elmhurst College, Elmhurst, IL 60126

In response to the publication of Vision and Change, the biology department at Elmhurst College revised our curriculum to better prepare students for a career in science with the addition of various writing assignments in every course. One commonality among all of the assignments is the ability to comprehend and critically evaluate scientific literature to determine relevancy and possible future research. Several previous reports have analyzed specific methodologies to improve student comprehension of scientific writing and critical thinking skills, yet none of these examined student growth over an undergraduate career. In this study, we hypothesized upper-level students would be better able to comprehend and critically analyze scientific literature than introductory biology majors. Biology students enrolled in an introductory (200-level), mid- (300-level), or late-career (400-level) course were tasked with reading and responding to questions regarding a common scientific article and rating their comfort and confidence in reading published literature. As predicted, upper-level (mid- and late-career) students showed increases in comprehension and critical analysis relative to their first-year peers. Interestingly, we observed that upper-level students read articles differently than introductory students, leading to significant gains in understanding and confidence. However, the observed gains were modest overall, indicating that further pedagogical change is necessary to improve student skills and confidence in reading scientific articles while fulfilling the Vision and Change recommendations.

INTRODUCTION

The publication of Vision and Change has led undergraduate biology departments to revise their curricula in order to arm graduates with the skills needed for future success in the sciences (1). To that end, our recently revised curriculum requires student preparation of various types of scientific writing in every biology course, incorporating either a learning-to-write or writing-to-learn mechanism in order to promote communication skills within the sciences (6, 13, 14, 16, 17).

While assessment of individual writing assignments and their ability to fulfill associated learning objectives is ongoing, a common aspect of all the writing assignments across the curriculum is the need to critically evaluate and gather information from the scientific literature. Previous studies have shown that reading the primary literature improves student comprehension and critical thinking skills, yet it is unknown how many undergraduate science departments are actively engaging students with the primary literature (3, 9, 12). To remedy this issue, recent studies have described different methods for teaching students how to read the literature (3, 8–11, 16, 19, 20). For example, Hoskins et al. described the C.R.E.A.T.E (Consider, Read, Elucidate hypotheses, Analyze and interpret data, Think of the next Experiment) method to approach primary literature reading (9, 10), in which considerable gains were observed in student critical thinking skills combined with the ability to integrate the scientific content (methods, results, and interpretation). However, this method is cumbersome, requiring a large amount of class time, or a dedicated course specific to reading the literature, that may not be available in a traditional undergraduate course or curriculum. Round and Campbell recently described their own technique, focused on a critical analysis of the data while “ignoring” aspects of the introduction and discussion (18). While this is likely a useful technique for students to critically assess the primary literature, it is also quite difficult for introductory students to assess figures of data when they do not have experience with experimental techniques. Recently, Segura-Totten and Dalman reported that the reading method does not matter and that all students who are given guidance in dissecting the primary literature gain critical thinking skills and show improvement in their ability to synthesize the significance of the reported findings (19). Science educators understand the importance of the primary literature to the field and the necessity for undergraduates to be comfortable...
comprehending and critically analyzing the literature. There is a dearth of research on the topic, however, and, in the case of the studies described above, none of them have examined changes in student comprehension and critical thinking skills over several years of an undergraduate career, focusing instead on one-course/one-term analyses (1, 7, 10, 15, 18).

In this study we examined the ability of both introductory and upper-level biology majors to critically evaluate a scientific article and demonstrate their ability to identify a hypothesis, determine major contributions to the field, and interpret graphics. We hypothesized that upper-level biology majors would be better able to successfully answer a series of questions regarding the chosen article than introductory biology majors. Here we report our findings about the ways in which students learn and evaluate the scientific literature and how that changes during the progression of the biology major.

**METHOD**

Participants and in-class assignment

Students in two sections of an introductory biology course for majors (BIO200 [General Biology I], n = 64) and two upper-level biology major courses (BIO355 [Evolution of Vertebrates], n = 18; BIO451 [Microbial Ecology], n = 17) were asked to read a journal article and complete a short questionnaire regarding the paper during the first week of the term. The upper-level courses differed slightly in the composition of students (Table 1). Of note, there were two students in BIO355 who had not yet completed the biology curriculum core and were taking an upper level course concurrently with Genetics; this was not the case for any BIO451 student. Also, there were twice as many transfer students (n = 8) in BIO355 as in BIO451 who had completed the introductory core courses, BIO200 and BIO201 (General Biology II), at another institution and had therefore not participated in the introductory learning-to-write assignments specific to those courses. Because the learning-to-write assignments include skill development in scientific research, reading, and writing, it is possible that transfer students may struggle with these skills in upper-level courses later in their careers. Finally, there was no significant difference in student grade point average between the students enrolled in BIO355 and BIO451.

Students were given thirty minutes during class to read Woese et al.’s article proposing a three-domain system of classifying organisms and answer a series of questions about the article (22; Appendix 1). The article was chosen for its significance, brevity (four pages), clarity and relative comprehensibility, a lack of highly detailed figures, and a concept to which students likely had no more than brief prior exposure. Students were given the article and questions together at the start of the 30-minute time period and could use the article when writing their responses to the assessment tool, but no additional instruction was provided. The first two questions focused on the students’ prior knowledge and their understanding of the article content to identify the hypothesis and key findings, both of which demonstrate lower-order cognition (2). Conversely, higher-level cognition was required for the third question, which assessed the ability of students to critically analyze and evaluate the paper to determine its significance within the greater scientific community (2). Finally, the remaining questions addressed students’ comfort and confidence with reading and analyzing scientific papers.

**Assessment methods**

Students’ responses to questions regarding the article content were scored as “excellent,” “good,” “adequate,” “incorrect,” or “unanswered.” In all cases, “excellent” responses were scored the highest with each subsequent scoring receiving one less point, respectively (Table 2). Student comfort and/or confidence in reading journal articles was self-reported on a scale of one to five, with a five indicating very comfortable or confident and a one representing not comfortable or confident (Appendix 1). Finally, to assess student reading methods, the number of times a student marked a journal article via highlighting or writing was recorded, specifically examining the type and location of marks within the journal article. To maintain consistency and objectivity, all surveys were blind scored by the primary author at the conclusion of the term using a previously agreed upon rubric, ensuring that student identity, upper- or lower-classmen standing, and participating course were unknown. All statistical analyses were conducted using SigmaPlot 12 using a Student’s t-test (Systat Software, Inc.).

**TABLE 1. Summary of student characteristics for each upper level biology elective.**

| Course                  | BIO355 | BIO451 |
|-------------------------|--------|--------|
| Mean Grade Point Average (GPA) \(^a\) | 3.374  | 3.437  |
| Number of transfer students in course | 8      | 4      |
| Number of students who had not finished core \(^b\) | 2      | 0      |
| Number of students taking first upper-level elective | 5      | 1      |
| Number of students taking second upper-level elective | 3      | 8      |
| Number of students with 3 or more upper-level electives | 10     | 4      |

\(^a\) There is no significant difference in GPA between upper-level electives.

\(^b\) These students were enrolled concurrently in the upper-level elective and the final biology core course, Genetics. Genetics is not a prerequisite for BIO355, Evolution of Vertebrates.
Informed consent and institutional review board protocols

Students signed an informed consent form prior to participating in this exercise and all articles and surveys were anonymous. Approval to conduct this study was granted by the Elmhurst College Institutional Review Board, which determined that the protocol fulfilled the necessary requirements for human subject research.

RESULTS

Assessment of student comprehension and critical analysis

Student comprehension and critical analysis of the journal article content were assessed with a three-question survey that was scored against a rubric and assigned a numerical result (Table 2). Out of a possible score of three, when asked to identify the hypothesis, a simple comprehension question, BIO200 students earned a mean score of 0.52 (17%) compared with 1.17 (39%) and 1.06 (35%) for BIO355 and BIO451 students, respectively (Fig. 1). Similarly when students were asked to identify the key findings presented in the paper, another question to examine their comprehension skills, BIO200 students had a mean score of 0.29/5 (5.8%) while BIO355 students scored 1.33/5 (26.6%) and BIO451 students scored 1.06/5 (21.2%). Finally, the third question focused on the ability of students to critically analyze the text and identify the contributions to the discipline. BIO200 students had a mean score of 0.4/2 (20%) while BIO355 students scored 0.83/2 (41.5%) and BIO451 students scored 1/2 (50%). In total, BIO200 students earned a mean score of 1.21/10 (12.1%) on the assignment while BIO355 students earned 3.33/10 (33.3%) and BIO451 students had a mean of 3.12/10 (31.2%). Statistical analysis using a Student’s t-test revealed that, in each case, there was no difference in scores between BIO355 and BIO451 students but there

\[ P < 0.001 \]

\[ P < 0.01 \]

\[ P < 0.005 \]

\[ P < 0.008 \]
was a significant difference between BIO200 students and the upper-level courses on every question (p values shown in Fig. 1).

**Student self-assessment**

Students were asked to self-report their comfort and confidence in reading the chosen article. When asked to identify the components of this journal article with which they struggled, the majority of students at all levels (81% for BIO200, 55.6% for BIO355, and 76.5% for BIO451) reported that they had difficulties understanding the language and wording of the article (Fig. 2). Most of the students (94%) in all groups surveyed felt that they understood the visuals in the journal article. Finally, 48% of BIO200 students compared with 11% of BIO355 and 0% of BIO451 students admitted to struggling with this article or the topic in general.

Students were also asked to rank, on a scale from one to five, their confidence with various aspects of reading journal articles in general. Using the indicated scale, both introductory and upper level students scored themselves high on their comfort and confidence in reading graphs and figures (3.59/5 for BIO200, 3.67/5 for BIO355, and 4/5 for BIO451), similar to what was self-reported on the previous survey question (Fig. 3). Students also rated their comfort with identifying the hypothesis of a paper high, with BIO200 students reporting a mean rating of 3.21/5 compared with a mean of 4/5 for BIO355 (p = 0.008) and 3.53/5 for BIO451 (no significant difference compared with BIO200). Ratings reflected lower confidence levels when it came to reading journal articles in general and understanding the methods employed. BIO200 students rated their confidence for these categories at 2.9/5 for each, while upper level students had mean ratings of approximately 3.4/5 for the same categories (p values shown on graph, Fig. 3). For both the introductory and upper-level student groups, the least confidence was reported in the ability to understand the problems or pitfalls within a particular paper, with only a 2.38/5 rating for BIO200, 3.33/5 rating for BIO355, and 2.85/5 for BIO451. Although there was a statistical difference in the ratings of BIO200 and BIO355 students in this area (p = 0.002), it was clear that all of the participants were least comfortable with criticizing published works.

**Examination of how students read journal articles**

Finally, after all materials had been collected and the assessment had concluded, it was observed that some students had marked up the provided journal article. No prior instruction had been given as to whether or not students should mark the article. Therefore, we analyzed these markings as a de facto measure of overall student comprehension. In some cases, these markings amounted to nothing more than a word or two underlined, while in other cases, many sentences were bracketed, underlined, or highlighted. When all photocopies were examined for each participating course, it was noted that upper-level students tended to make many marks on their copies of the paper, while students in introductory classes made noticeably fewer marks on their photocopies (Fig. 4a). Overall, only 31% of students in the introductory course wrote on the paper provided to them, while 67–70% of upper level students made comments or markings on their paper.

When the number and type of markups were examined and compared between all three participating courses, it was noted that introductory students typically made only...
limited markings on their photocopied article (Fig. 4b). In fact, of the 20 introductory students who made marks on their paper, 12 were limited to only one to five markings. Further examination of these papers revealed that five of these students highlighted one or two sentences in the abstract with no further markings on the paper and, in each case, the words highlighted included the three proposed domain names (Archaea, Bacteria, and Eucarya; Table 3). Regardless of the degree of marking or highlighting on a paper copy, a recurrent observation in all of the marked copies of the papers used by the introductory students was that students tended to mark/underline either the outdated kingdom names or the newly proposed domain names (data not shown). Furthermore, introductory students were far more likely to highlight their papers primarily in the abstract and first section of the paper, but did not mark as much in later sections where the actual findings of the paper, implications, and conclusions were found (Table 3). Conversely, participating upper-level biology students tended to highlight or underline much more often on their photocopies. Nearly 43% of upper level students surveyed marked up their papers on 11 or more sentences compared with the 6% of introductory level students (Fig. 4b). At the same time, upper-level students were more likely to highlight the paper throughout all sections rather than focusing on the abstract and introductory material (Table 3). Interestingly, very few students wrote comments on the paper itself. Instead, the majority of all markings on the papers were in the form of highlighting of information that students deemed important. Only two students in the BIO200 course (out of 64) and three students in BIO451 (out of 17) wrote on the article itself (Table 3).

DISCUSSION

The ability to comprehend and critically evaluate scientific literature demonstrates not only students’ critical thinking skills but also their ability to think like scientists to identify gaps in the literature and ask questions for future research. Indeed, these are the types of skills that are imperative in training the next generation of scientists and it is an important aspect of the Vision and Change model (1). While there have been multiple studies describing effective methods to develop critical reading skills, none of these have examined continued development of reading and comprehension skills over time. Furthermore, it is well documented that STEM students, in particular, thrive in active learning environments, and analyzing the scientific literature is a tool in that process (4). Here, we have developed a way to assess how effective we have been at giving students the tools they need to prepare for and be successful at active learning.

Our overall findings indicate that upper-level students show significant gains in their comprehension of the scientific literature compared with introductory students. However, their overall scores (3.3/10 and 3.1/10, Fig. 1) indicate that there is still much room for future learning. These scores indicate an increase in understanding of the topic for this cohort of upper-level students compared with introductory students but not to the level that we hope our students will eventually achieve. Furthermore, the gains shown in previous studies (9, 19) are also seen in our work. However, while most of the gains observed here are statistically significant, they are still modest. Particularly concerning is the lack of comments written on the article by students of any level (Table 3). Written comments would indicate students are engaging in active learning and critical analysis. Merely highlighting the article, even when comprehensive, suggests passive learning.

Based on these modest gains, this study gives us an idea of where pedagogical changes are needed in the curriculum.
Currently, no faculty in the department use a specific method (e.g., C.R.E.A.T.E.) to teach students how to read the literature. Instead, each individual faculty member instructs students in a more generalized format without an underlying pedagogical rationalization, while possibly overemphasizing scientific content. Based on these data, a more interventionist approach could be warranted, and we would expect that a more formalized method of instruction could lead to greater gains in reading comprehension and critical analysis in the future based on previous findings with C.R.E.A.T.E. and other methods (9, 19).

When determining how to utilize a method such as C.R.E.A.T.E., it will have to be applied equally across all upper-level electives, perhaps as part of further curriculum revisions. Though there were twice as many first-time upper-level students in BIO355 as in BIO451, there were also more experienced students taking their fourth or fifth elective in BIO355 than in BIO451. This is due to unrestricted enrollment and results in the unpredictable composition of our upper-level electives, which was reflected in these data as a lack of significant difference in the performance of students in the two upper-level courses. Formative reading comprehension instruction could begin at the BIO200 level, advance in BIO315, Genetics, and be completed in the upper-level electives. Reading comprehension exercises would nicely complement our developmental learning-to-write and writing-to-learn assignments (14).

One of the more surprising findings in our study was how students actually read and mark up an article. As stated previously, we did not provide any instruction to the students on whether they could, or should, mark the article. Upon receiving the anonymous responses, the differences in reading styles between introductory and upper-level students was remarkable. It was clear that upper-level students felt much more comfortable marking up an article than introductory students. And, while the increase in highlighted areas on the photocopy does not indicate the degree to which students understood the content, it may correlate to how students are learning to read journal articles. Again, as no formal methodology is taught in any course within the department, it is interesting to note which skills may be self-taught or perhaps learned in courses outside of the sciences in which students read many papers/ texts. Furthermore, the type and degree of markings was clearly different between the two groups of students, with upper-level students making more marks throughout the paper while introductory students often just highlighted key words within the abstract. This difference can be seen in the scores of introductory and upper-level students when asked to determine the key findings of the paper. BIO200 students scored 5.8% on this question compared with a nearly fourfold increase of 26% and 21%, respectively, for BIO355 and BIO451 (Fig. 1). While a 21% score does not indicate true comprehension of the topic, the difference between BIO200 and upper-level student scores is remarkable. Based on the fact that the majority of introductory students only marked their journal articles in the first section (Table 3), prior to the sections of the paper that include the key findings, these results may not be surprising. Instead, it may indicate that introductory students do not truly understand where the important and relevant information in a journal article would be found and, instead, remain focused on the background material, which they recognize. This raised the question of whether the introductory students were unable to comprehend the article itself and were only marking words that they had previously heard of in other science courses. On the other hand, we also note that upper-level students were also more likely to make more than 20 markings on the article, leading to an article that was covered with markings, making it difficult to distinguish the most important material. Obviously, extremes on either end are ineffective methods for journal reading and indicate that there is still work to be done with all levels of students.

### Table 3.
Breakdown of location and type of student markups on journal articles.

| Location of Markings                  | Percent of Total Markingsa |
|--------------------------------------|----------------------------|
|                                      | BIO200b | BIO355c | BIO451d |
| Abstract                             |     24.2 |     8.4  |    13.3  |
| Need for Restructuring Systematics   |     43.3 |     33.0 |     37.0 |
| Basis for Restructuring              |     18.3 |     34.6 |     22.0 |
| Proposal for a New Highest Level Taxon|     11.7 |     21.8 |     23.1 |
| Definitions                          |       0  |       0  |       0  |
| Conclusion                           |     2.5  |     2.2  |     4.6  |

---

*a Total number of markings was 120 for BIO200, 179 for BIO355, and 173 for BIO451.
*b Two papers had one comment written in the margins: 1 in the abstract and 1 in the Need for Restructuring Systematics.
*c No comments were written in the margins by BIO355 students.
*d Three papers had comments written in the margins: (1) 1 in the Proposal for a New Highest Level Taxon; (2) 2 in the Need for Restructuring Systematics, 1 in the Basis for Restructuring, and 2 in the Proposal for a New Highest Level Taxon; and (3) 2 in the Need for Restructuring Systematics.
Finally, another interesting finding from our study—though perhaps not surprising—was students’ self-assessment of their confidence and comfort in reading journal articles compared with their actual ability. One of the most striking examples of this was in the area of comprehension. While introductory students appeared to be more honest in admitting that they did not comprehend the article well, which was shown in their overall scores, almost half of all upper-level students indicated they were confident in their comprehension of the material, even though their average score was a 3.4/10, well below the 50% mark. These data indicate a disconnect between students’ opinions of their ability to read articles and their actual skill level. This disconnect has been shown to be lessened by inquiry-based assignments such as working in groups on problem solving, using personal response systems, such as clickers, and completing in-class case studies (4, 5). While the results of this assessment were kept anonymous for publication purposes, it may be useful in the future to use this assignment as a teaching tool for students in order to show the disconnect and remind them that there is still work to be done. Improving students’ abilities to more accurately assess their own skill levels could have an impact beyond the writing assignments and aid students in developing metacognitive skills (21). Furthermore, these findings may indicate a need to incorporate more formalized journal discussion and/or write-ups within courses in the major to give students additional experience in this area and see how they compare with their peers in a classroom setting. Currently only a few select courses in the major utilize journal discussions, and several of these courses may only discuss one or two articles throughout the entire semester, which may not provide enough time for student self-assessment, feedback, and growth. As we continue to assess our curriculum and critical thinking skills within the curriculum, these data may inform future pedagogical changes in individual courses.

In conclusion, the biology department of Elmhurst College revised its curriculum in a way that conforms to Vision and Change initiatives. Assessment of individual components of the curriculum is essential to the successful implementation of Vision and Change directives. This initial assessment of more passive student skills, such as literature comprehension, establishes the groundwork for future assessments of the active learning components of the curriculum, such as original research. This analysis also suggests that even passive skills such as reading comprehension can be transformed to incorporate more active learning and critical analysis. Any initial changes to our curriculum must be guided by efforts to improve critical thinking skills before attempting to modify the way in which students apply those skills within our curriculum. This early assessment and the improvement of student scores confirms that the cohesive design of the curriculum is succeeding in building upon skills introduced in the biology core courses.

SUPPLEMENTAL MATERIALS

Appendix I: Journal article survey

ACKNOWLEDGMENTS

The authors declare that there are no conflicts of interest.

REFERENCES

1. American Association for the Advancement of Science. 2011. Vision and change in undergraduate biology education: a call to action: a summary of recommendations made at a national conference organized by the American Association for the Advancement of Science, July 15–17, 2009. Washington, DC.
2. Bloom, B. S., and D. R. Krathwohl. 1956. Taxonomy of educational objectives: the classification of educational goals, by a Committee of College and University Examiners, Handbook I: Cognitive Domain. Longman, New York, NY.
3. DebBurman, S. K. 2002. Learning how scientists work: experiential research projects to promote cell biology learning and scientific process skills. CBE Life Sci. Educ. 1:154–172.
4. Freeman, S., et al. 2014. Active learning increases student performance in science, engineering, and mathematics. PNAS. 111(23):8410–8415.
5. Gormally, C., P. Brickman, B. Haller, and N. Armstrong. 2009. Effects of inquiry-based learning on students’ science literacy skills and confidence. Int'l. J. Scholarship Teach. Learn. 3(2):1–22.
6. Graham, S., and D. Perin. 2007. Writing next: effective strategies to improve writing of adolescents in middle and high schools. Alliance for Excellent Education, Washington, DC.
7. Handelsman, J., et al. 2004. Scientific teaching. Science 304:521–522.
8. Hoskins, S. G. 2008. Using a paradigm shift to teach neurobiology and the nature of science—a C.R.E.A.T.E.-based approach. J. Undergrad. Neurosci. Educ. 6:A40–A52.
9. Hoskins, S. G., D. Lopatto, and L. M. Stevens. 2011. The C.R.E.A.T.E approach to primary literature shifts undergraduates’ self-assessed ability to read and analyze journal articles, attitudes about science and epistemological beliefs. CBE Life Sci. Educ. 10:368–378.
10. Hoskins, S. G., L. M. Stevens, and R. H. Nehm. 2007. Selective use of the primary literature transforms the classroom into a virtual laboratory. Genetics. 176:1381–1389.
11. Jacques-Fricke, B., A. Hubert, and S. Miller. 2009. A versatile module to improve understanding of scientific literature through peer instruction. J. Coll. Sci. Teach. 39:24–32.
12. Kolzeracki, C. A., M. F. Carey, J. Colicelli, and M. Levis-Fitzgerald. 2006. An intensive primary-literature-based teaching program directly benefits undergraduate science majors and facilitates their transition to doctoral programs. CBE Life Sci. Educ. 5:340–347.
13. Moskovitz, C., and D. Kellogg. 2011. Inquiry-based writing in the laboratory course. Science 332:919–920.
14. Raimondi, S. L., T. L. Marsh, and P. E. Arriola. 2014. Integrating Vision and Change into a biology curriculum at a small comprehensive college. J. College Sci. Teach. 43(5):33–39.
15. Ramos Goyette, S., and J. DeLuca. 2007. A semester-long student-directed research project involving enzyme immunoassay: appropriate for immunology, endocrinology, or neuroscience courses. CBE Life Sci. Educ. 6:332–342.
16. Reynolds, J. A., C. Thaiss, W. Katkin, and R. J. Thompson, Jr. 2012. Writing-to-learn in undergraduate science education: A community-based, conceptually driven approach. CBE Life Sci. Educ. 11(1):17–25.
17. Rivard, L. P. 1994. A review of writing to learn in science: implications for practice and research. J. Res. Sci. Teach. 31:969–983.
18. Round, J. E., and A. M. Campbell. 2012. Figure facts: encouraging undergraduates to take a data-centered approach to reading primary literature. CBE Life Sci. Educ. 12(1):39–46.
19. Segura-Totten, M., and N. E. Dalman. 2013. The CREATE method does not result in greater gains in critical thinking than a more traditional method of analyzing the primary literature. J. Microbiol. Biol. Educ. 14(2):166–175.
20. Switzer, P. V., and W. M. Shriner. 2000. Mimicking the scientific process in the upper-division laboratory. BioScience 50:157–162.
21. Tanner, K. D. 2012. Promoting student metacognition. CBE Life Sci. Educ. 11:113–120.
22. Woese, C. R., O. Kandler, and M. L. Wheelis. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. 87:4576–4579.