Article

Concepts First, Jargon Second Improves Student Articulation of Understanding[S]

Lisa McDonnell†*
Megan K. Barker‡
Carl Wieman§

From the †Division of Biological Sciences, University of California San Diego, San Diego, California, ‡Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada, §Department of Physics and Graduate School of Education, Stanford University, Stanford, California

Abstract

In this experiment, students in a large undergraduate biology course were first exposed to the concepts without new technical vocabulary ("jargon") in a pre-class reading assignment. Their learning of the concepts and jargon was compared with that of an equivalent group of students in another section of the same course, whose pre-class reading presented both the jargon and concepts together in the traditional manner. Both groups had the same active-learning classes with the same instructor, and then completed the same post-test. Although the two groups performed the same on the multiple choice questions of the post-test, the group exposed to concepts first and jargon second included 1.5 times and 2.5 times more correct arguments on two free-response questions about the concepts. The correct use of jargon between the two groups was similar, with the exception of one jargon term that the control group used more often. These results suggest that modest instructional changes whereby new concepts are introduced in a concepts-first, jargon-second manner can increase student learning, as demonstrated by their ability to articulate their understanding of new concepts.

Keywords: *vocabulary; learning; concept; introductory biology*

Introduction

Scientific literacy, defined as "the knowledge and understanding of scientific concepts and processes" (United States National Center for Education Statistics), is a central goal for many undergraduate science programs. In developing scientific literacy within a specific discipline, it is necessary to gain fluency with the fundamental concepts, and the technical vocabulary used to describe these concepts. One study has shown that there are more new terms in science textbooks than in foreign language classes [1]. The meaning of much of the technical vocabulary terms used in science is not always intuitive to a novice, and hence it becomes "jargon." The problem of teaching jargon-heavy concepts is widely known anecdotally among instructors, and has also been identified in the literature as a potential barrier to learning science [2–7]. Particularly in biology, this "vocabulary load" may negatively impact student learning [1, 3, 8].

To address the vocabulary problem, one proposed teaching strategy is to reduce the number of concepts or new terms introduced in a textbook or course [3, 9]. Another is to develop additional outside-of-class activities to support students' learning of vocabulary [10]. While these approaches certainly have value, increasing student workload with out-of-class interventions and/or implementing curricular changes are often outside the control of individual instructors. Given the reality of curriculum constraints on many introductory courses, we sought a different, practical, and scientifically testable approach toward reducing the jargon barrier in one biology topic.

To our knowledge, there has been no experimental study that specifically targeted the impact of jargon on conceptual understanding in undergraduate biology. One related study used a comprehensive active learning approach in second-year genetics, with language emphasis including the presence of a language expert in lectures and tutorials who guided interventions in these sessions [11]. However, they did not measure

---

Volume 44, Number 1, January/February 2016, Pages 12–19

[S] Additional Supporting Information may be found in the online version of this article.

L.M. and M.B. contributed equally to this work.

**Address for correspondence to: Division of Biological Sciences, University of California San Diego, San Diego, California, USA. E-mail: lmcdonnell@ucsd.edu.*

Received 3 March 2015; Revised 28 August 2015; Accepted 9 September 2015

DOI 10.1002/bmb.20922

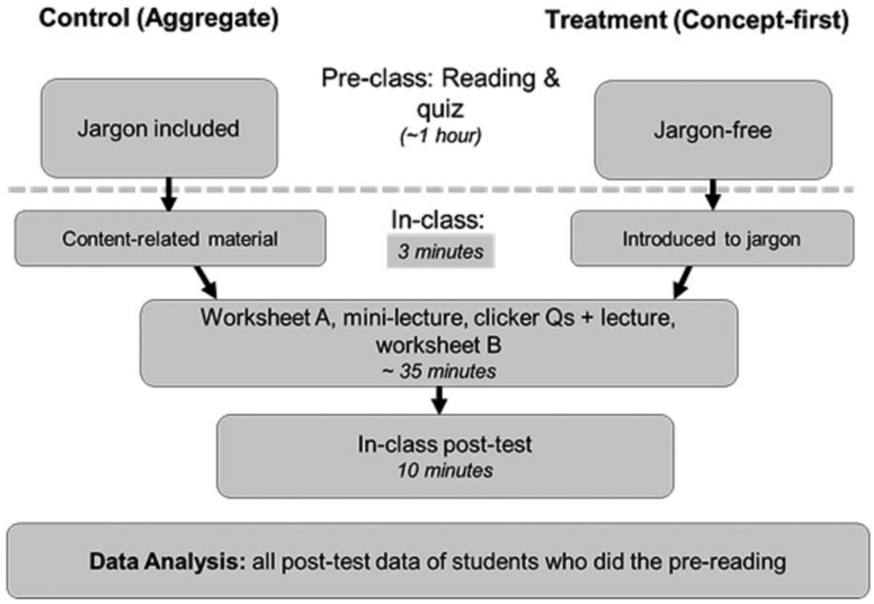

FIG 1

A flowchart of the experimental design. The experimental treatment (concepts-first) largely took place outside of class time, in the form of a pre-class reading and quiz. Both the control and concepts-first groups had the same reading and quiz, with the exception of the jargon being replaced as per Table I in the concepts-first materials. In-class, students in the treatment group were briefly introduced to the jargon by reading, while the control group received a few minutes of reading content-related material. Subsequently, the in-class activities and post-test were identical in the control and treatment groups (clicker Qs are questions posed and answered using the iClicker personal response system). The post-test is included in the Supporting Information.

any changes in overall performance relative to prior years. Thus, the broad use of various active learning approaches may not be sufficient to address the specific challenge of learning technical vocabulary in undergraduate science.

At the elementary school level, Brown and Ryoo [8] measured greater learning gains on end-of-unit tests when jargon was removed from the initial learning phase of a new topic (photosynthesis). Inspired by this work, we sought to investigate the relationship between the introduction of jargon and student conceptual learning of a new topic in undergraduate biology. We hypothesized that initially replacing jargon with more familiar terms would improve student learning. Our hypothesis is founded on the cognitive load theory [12], which suggests that jargon increases the cognitive load and could therefore decrease learning. Testing this hypothesis within two undergraduate biology sections in the context of the "DNA structure and the genome" topic, we disaggregated students' initial exposure to new jargon and new concepts, and assessed student learning on a post-test. Our results showed significant improvement in students' articulation of the relevant concepts when new jargon and concepts were separately introduced. Thus, we propose an instructional approach, one which can be done within a regular undergraduate lecture, to improve student learning of concepts.

Methods

Course Background and Study Design

Data collection for this study took place during the Winter 2014 term of a first-year introductory cell biology course at the University of British Columbia. There were two lecture sections of approximately 230 students each; a subset ($n = 42$ students in each section) was included in data analysis based on completion of all the experimental materials. The study design was similar to that of Brown and Ryoo [8], whereby the control group was introduced to new concepts and jargon simultaneously, while the treatment group ("concepts-first") was introduced to the same concepts but with the jargon replaced with everyday language. Figure 1 summarizes the study design.

Pre-Class and Classroom Activities

The experiment followed normal course structure for this class: all students were assigned pre-class reading (accessed online) followed by completing a short graded quiz (online), and then attending a 50-min lecture. Pre-reading for the control group consisted of a short section of the textbook on the material to be covered in the upcoming class, while the concepts-first group's pre-reading included the same passage and figures, presented in the same order, except the specifically chosen jargon terms were replaced with everyday language (Table I, and Supporting Information, Table S1 for a sample of the reading). That group was then introduced to the jargon at the beginning of the class. The two lecture sections are normally each taught by two different instructors; for this study, one of the authors (MKB) delivered the lectures to both the control and concepts-first sections, while another author (LMM) attended/timed them to monitor consistency. The lecture

TABLE I *Jargon and substitute terms identified for this study and assessed on the post-test*

Learning objective/topic	*Jargon term*	*Substitute term*
Explain the effect of a mutation on the structure and stability of DNA	Purine	Large base
	Pyrimidine	Small base
	Stacking interaction	Hydrophobic interaction
Identify and explain what a genome is	Genome	Total hereditary genetic material

style was consistent with the general style of the course, which implements an active-learning strategy with in-class activities designed to support student learning [13], including grades for participation. The distribution of class time was the same in both sections (Fig. 1).

Content Topics and Jargon Substitution

The topic selected for this study was introductory DNA structure and the genome. This was chosen based on the amount of technical vocabulary normally included in this course unit, as well as considerations to minimize the amount of student's prior knowledge on the subject. Within the course, this topic begins a new unit and is typically covered in the first lecture following a course midterm, so there is minimal prior exposure to it. Jargon within this topic was identified by the course instructors and study authors who had identified vocabulary in this unit as problematic for students. The terms identified as jargon, and substituted with plain language, and assessed in this experiment, are indicated in Table I. The substitute terms/phrases were chosen by the LMM and MKB, in consultation with the course instructors; to capture the most relevant information students were required to understand in this unit from the scientific term, in plain language, or language that had been used previously in the course for similar phenomena.

Study Cohort

The sizes of the participant groups included in this study are presented in Table II. To ensure that we were measuring the effect of the jargon replacement in the pre-reading, we established cohorts of students in both sections who reported that they fully completed the pre-reading (clicker question asked in class; see question in Supporting Information, Fig. S1). Our primary comparisons of post-test results included only those students who selected "I read all of the pre-reading before today's pre-quiz". While this substantially decreased the number of students in the study, we took this conservative approach to ensure that all students in the study had fully experienced the experimental variable. Students who selected "I didn't read the pre-reading for today" were used for additional comparisons as discussed below. All students had taken a common midterm 1 week prior to this experiment, and the scores on the midterm were used to compare the control and treatment group populations (Table II). The students in the control and treatment groups who did the pre-reading were equivalent based on pre-experiment midterm scores (t-tests $p > 0.1$), although the non-readers from the control group performed slightly, but not significantly (t-test $p > 0.1$) lower on the midterm.

Assessment and Analysis of Student Performance

Student performance was measured on an in-class post-test completed individually. The test assessed two specific topics within the course material: (1) the chemical interactions that stabilize DNA structure and (2) the information content of a genome. The post-test consisted of two pairs of multiple-choice questions (one pair per topic), and two short free-response questions (one per topic). Each multiple choice pair was isomorphic, including one question that included jargon, and one that did not. Using midterm scores to sort students into quartiles, the discrimination index of each multiple choice question was determined. Each question had a very similar discrimination index (Q1 0.22, Q2 0.21, Q3 0.18, Q4 0.26). The free-response questions did not include jargon in the prompt; all questions can be found in the Supporting Information. Multiple choice questions were administered by projecting the question on the slide. Students were given a set amount of time for each question (approximately 1 min) and the remaining time was devoted to the free-response questions (approximately 6 min). All questions were reviewed by a second course instructor (not involved in the research and blind to the project goals) to confirm the questions were clear, lacking jargon (where appropriate), and testing the desired concepts.

All post-tests were analyzed blind to experimental condition (treatment or control). Multiple choice questions were analyzed for correctness (only one correct answer per question). Free-response questions were scored for correct use of jargon and the total number of correct arguments

TABLE II *Class size and participant information*

Information	*Control*	*Concepts-first*
Number who participated in the class, as measured by numbers of students who wrote the post-test at the end of class	229	231
Number who completed the experimental pre-class reading assignment (and were subsequently included in analysis)	42	42[a]
Mean midterm score of students who did the pre-class reading (standard deviation)	74% (12%)	77 (14%)
Number of students who self-reported to have not completed the full pre-reading	22	21
Mean midterm score of students who did not do the pre-class reading (standard deviation)	68% (12%)	74% (14%)

[a] Seventy-seven students completed the pre-class reading, but 42 were randomly selected from these 77 to keep the control and treatment group sizes consistent for analysis.

included in the answer (see rubric in Supporting Information, Table S3). The rubric/criteria for determining if an argument was correct emerged using a process derived from grounded theory [14], involving an iterative process of blind reviewing student responses. For the first iteration, each reviewer (MKB and LMM) independently reviewed a subset of randomly selected student responses, documented the correct arguments included in student responses, and identified arguments that met the requirements of the question. Comparing these reviews led to development of a rubric to allow for scoring of multiple correct explanations (see Supporting Information, Table S3). MKB and LMM then used this rubric to score the students that met the full-reading criteria ($n = 42$ in each section). Comparison of MKB and LMM scoring revealed greater than 95% inter-rater reliability, and any differences were resolved through conversation.

Results

The most striking difference between control and concepts-first groups is the number of correct arguments included in answers to the free-response questions (Fig. 2). The concepts first group provided 2.5- and 1.5-fold more correct arguments than that of the control group, on the DNA structure and genome topics, respectively. The breakdown of the number of correct arguments is given in Table III; the significant difference is due to many more students in the concepts-first group having one or two correct arguments in their answers to the free-response questions, compared to the majority of students in the control group with no or one correct argument.

A finer-grained view of the students' conceptual understanding can be seen in Fig. 3. The DNA structure question had two conceptual components considered correct and relevant to the question: (1) that the specified mutation causes a physical change in the size of the base-pair, which will alter the interstrand distance and (2) an alteration of the interactions that stabilize the structure. The concepts-first group included at least one correct argument about concept 1 more commonly ($n = 22$ students concepts-first, $n = 13$ control, chi-squared test $p < 0.01$) while the two

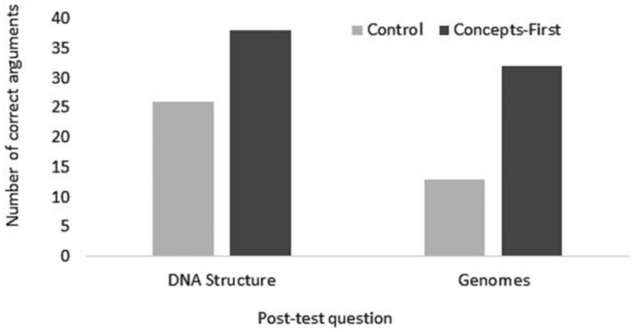

FIG 2 *Total number of correct arguments on the free-response questions. Relative to the control population, a significantly larger number of correct arguments were measured in the written answers of students in the concepts-first group for both the DNA structure and genomes post-test questions.* n = *42 students for each of the control and concepts-first groups. Note that the total possible number of correct arguments provided by a group of 42 students could be 126 (according to the rubric), indicating that these were challenging questions for the students.*

TABLE III *The number of student responses that included 0, 1, 2, or 3 correct arguments on the free-response post-test items*

Number of correct arguments	DNA structure		Genomes[a]	
	Control	Concepts-first	Control	Concepts-first
0	26	18	31	22
1	8	15	9	9
2	6	4	2	10
3	2	5	0	1

There were three possible arguments for each topic/question (see rubric in Supporting Information).

[a] *indicates statistical significance at $p < 0.05$ on a Fisher's exact test comparing control and concepts-first within each question topic.*

groups showed no difference in their inclusion of a correct argument about concept 2 ($n = 25$ students concepts-first, $n = 23$ control, chi-squared test $p > 0.5$). The same trend was measured for the genome question (Fig. 3), which had two conceptual components: the genome consists of all the (haploid) genetic material (coding and noncoding DNA sequences) in a cell, and it is hereditary. Students from the concepts-first group provided more correct arguments in their answer to this question than did the control group (Table III, Fisher's exact, $p < 0.05$). In fact, no students from the control group included an argument about the hereditary nature of the genome, whereas 21% of the concepts-first group included such an argument (Fig. 3).

The correct use of the jargon terms "stacking interactions," "purine," and "pyrimidine," and "genome" were scored in student answers to the free-response questions. The mean percent score for correct use of jargon terms in student responses was low, with a nonsignificant trend in favor of the control group (30% and 25% correct for control and concepts-first, respectively). The slightly higher correct use of jargon by the control group was due to differences in correct use of "stacking interactions" in their responses compared to the treatment group (36% treatment, 21% control; chi-square $p = 0.05$). The other two jargon terms were used correctly with equal frequency by students in the control and concepts-first groups (purine/pyrimidine used correctly by 21% of students in both groups; genome used correctly by 33% of students in both groups).

The post-test also contained four multiple choice questions: two with jargon and two without jargon. There were no significant differences in overall scores or the percentage of students correct on a question-by-question basis between control and concepts-first groups (t-test $p > 0.05$ for all questions; average scores shown in Supporting Information, Fig. S2).

Analysis of Cohorts Who Did Not Complete the Reading

The only differences between the control and treatment conditions were (1) the pre-reading treatment and (2) the first 3 min of class, in which vocabulary was introduced (treatment) or content-related material was presented (control). To account for any effects of the second factor, we analyzed the scores of students who did *not* complete the pre-reading. In the concepts-first group, students who reported that they did not complete the pre-class reading performed significantly worse on both the free-response and multiple choice questions (Supporting Information, Fig. S3 and Table S2). The same trend was observed in the control group, while the only significant difference was observed in the free-response scores and not multiple choice scores. These findings indicate that the pre-reading was beneficial (or that stronger students do the pre-reading in the first place), and, more crucially to our study, that there was no major learning difference imparted by the small portion of class where vocabulary was introduced.

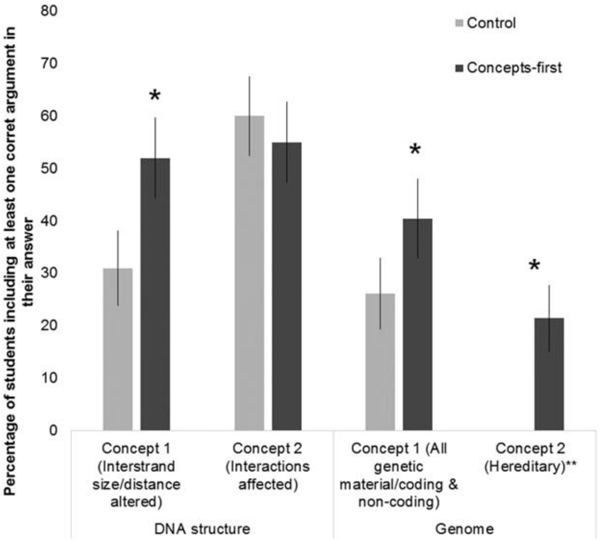

FIG 3 *Students' correct arguments on specific open-response questions. Statistically significant differences are seen between concepts-first and control for two of the three arguments used. Details of the arguments can be found in the text, and in the rubric provided within the Supporting Information. * indicates statistical significance at $p < 0.05$ on a Chi-squared test (**Genome concept two had a 0 result for the control group; hence, a Fisher's exact for Genome concept two because of the 0 result in the control group) comparing the number of student responses with said correct argument between concepts-first and control. Error bars are standard error of the mean for binomial data.*

Discussion

The aim of this research was to determine the impact on student learning of presenting new material to students in a concepts-first and jargon-second approach. We hypothesized that substituting jargon with everyday language would improve learning of the concept (as demonstrated by improved performance on a test). In support of this hypothesis, these results show that students who first saw a jargon-free explanation of the concepts performed better on the free-response questions, including more correct arguments in their answers.

The composition of written arguments indicates that jargon substitution had a positive impact on student understanding. For example, we substituted "purine" and "pyrimidine" with "large base" and "small base," respectively, to indicate the relationship between the base and the space it occupies in the DNA molecule. In response to a question asking about the effects of noncomplementary base-pairing on the stability of the DNA molecule, there was no difference between groups in the number of students that stated interactions within the DNA molecule would be affected by the noncomplementary base-pairing. However, significantly more students in the concepts-first group identified that the size of the base would affect the structure of the DNA molecule, or that interstrand distance would be affected by the bases. Our results suggest that students in the concepts-first condition acquired a better understanding of the relationships between base sizes and the structure of a DNA molecule, and were therefore able to predict and articulate the effects of the change to base-pairing more successfully than students in the control group. Likewise, student responses to the genome question included a larger number of correct arguments of what the genome is, indicating a better understanding of genome structure and content. Students in the concepts-first group who reported that they did not engage in the pre-reading performed significantly worse on the post-test, further supporting that it was the experimental treatment of jargon-free pre-class reading that had an impact on learning.

Our results are consistent with Brown and Ryoo's findings at the elementary school level [8]: student performance was higher when jargon was removed from the initial learning phase of a new topic. Those authors interpreted their results in terms of the effects of discursive identity on student learning (i.e., student identity as communicated through language). Both their and our findings may also be interpreted in light of cognitive load theory [12, 15, 16]. The large amounts of jargon that students are exposed to when being introduced to new concepts in biology classes may produce a large cognitive load that will negatively impact learning [17, 18]. One example of this is the overall reduced learning and performance when students were asked to learn a new skill and a mathematical concept within the same task, compared to when the concept and skill are taught separately [18]. The cognitive load of learning concepts and skills simultaneously is analogous to a typical biology class, whereby students are tasked with learning a new concept while simultaneously learning new jargon, make connections between the two, and integrating this knowledge and vocabulary into their larger, existing framework of understanding. We did not directly measure if jargon increased cognitive load during our study; such work would be an important future study to better understand the impact of jargon on learning.

In interpreting these findings, an additional consideration is the relationship between the key aspects of the concept and the application of that concept. For example, the generalization of purines/pyrimidines as large/small bases, rather than as some description of their other chemical properties, is particularly relevant to the application and the learning goal within this course—describing the impact of mutations on higher DNA structure. Learning about the concepts with simpler or more familiar terms may make the relevant concepts more accessible to students, and hence make it easier for them to integrate the new information into their existing knowledge frameworks and apply this new knowledge. Organization of knowledge, the number of connections between concepts and pieces of knowledge, impacts one's ability to retrieve information and apply knowledge to solve problems [19], which could be one explanation as to why more students in the concepts-first group were able to connect the concepts learned to solve the problem being asked. This highlights the importance in choosing a substitute term that is familiar but also focusses on the aspect of the concept that is most important in a given context/course; further work that investigates both a classification of initial jargon terms, and subsequent optimal selection of replacement terms, would be beneficial.

Students in the control and concepts-first groups performed similarly on the multiple-choice questions, irrespective of the presence/absence of jargon terms. This result was not surprising to us as the ability to recognize a correct statement with jargon is less cognitively challenging—at a lower level on Bloom's taxonomy—than having to synthesize the concepts and relationships to answer a question with a written argument [20]. Additionally, these results suggest that the jargon substitution terms and phrases we used were intuitive enough such that students in the control group were not at a disadvantage when jargon was not used, validating the choice of replacement terms. Success on the multiple choice test was not correlated with higher numbers of correct arguments on the free-response questions; thus in future, we would aim to develop more challenging multiple choice questions and rely more heavily on free-response answers as our indicator of conceptual understanding.

Overall performance on the post-test was unsurprisingly low, indicating that more time is required for students to master both the concepts and the jargon, and become more fluid at moving between the jargon and the concepts they represent. This is not surprising, as deliberate practice is required for mastery [21]. However, the fact that we saw any learning gains after such a modest instructional change, and after minimal student time interacting with the material, is quite a promising finding for educational impact.

Student Use of Jargon, and Types of Jargon as Barriers to Learning

In general, students' use of the jargon terms was quite low, suggesting that students likely need more time incorporating jargon into their framework of conceptual understanding, and more time practicing using jargon in written arguments. Students in the control group used the jargon term "stacking interactions" more frequently than the concepts-first students. This prompted us to question whether "stacking interaction" is truly jargon. This term was initially identified because students had not encountered this term before, whereas students had learned about hydrophobic interactions earlier in this course. However, if students do not have a firm understanding of what a hydrophobic interaction is, they will not have the framework to understand these interactions in the context of DNA structure, as a particular type of "stacking" interaction. In retrospect, based on these data, we believe that "stacking interaction" is a more accessible description of the hydrophobic interactions that occur within the DNA molecule. Thus, despite prior exposure and testing on the previous midterm, "hydrophobic interaction" is less understandable to the students, which may explain why the control group outperformed the treatment: in this case, the control group had been first exposed to the (new) plain language rather than the (old) jargon.

Study Limitations and Future Work

This reasoning brings to light a limitation of this study and an important area for future work: the analysis and selection of what is, and is not, considered jargon. Aside from emphasis placed by the textbook, and consensus from the instructors who have taught this subject before, we did not have any direct measures of student familiarity with vocabulary in this topic, and hence which terms did or did not pose a challenge when learning a new concept. However, as was the case with the term "hydrophobic interaction," students and instructors may not agree on which terms students are already fluent in, despite prior exposure and testing within the same course. Future experiments that first identify broad characteristics of discipline-specific jargon types, and then make connections to how they can help or hinder learning, would be highly valuable from a teaching standpoint. To begin addressing these issues, we are currently utilizing surveys to directly capture student understanding of and perceptions around difficult jargon.

Aside from an exploration of jargon types and using student feedback to select appropriate everyday-language substitute terms, more work is required to fully understand how changes in instructional design of jargon-laden topics may impact student learning. An area of future work would be to explore larger structural changes to reduce the negative impact of unfamiliar jargon: for example, modifying the in-class treatment of jargon, and testing these ideas on longer timescales. In line with our current findings, it is likely that these strategies will provide even greater improvements on student learning.

Implications for Teaching

STEM experts agree that students need to develop the ability to use specialized language in order to communicate their understanding of concepts. The problem associated with large amounts of jargon in biology is not a new idea, and the literature is scarce on practical approaches to address the issue within the context of a given undergraduate course. In our work, the significant differences between the control and treatment group were the result of a very small intervention. Changes were made in only a fraction of the time students were learning about this material, during a time without direct instructor contact—the pre-class reading for 1 week. We did not increase class time, reduce course material, or increase student workload. This instructional approach can be adapted to most disciplines, and could be useful for those that construct the reading material (textbook authors, or instructors who make their own pre-reading material). It is also reasonable to extend these findings for use in classrooms where students' first exposure to material is during lecture. Organizing the lectures to introduce concepts first and later include jargon is a small instructional change that will likely have positive impact.

Our results show that the substitution of jargon with everyday terms and phrases can significantly improve student understanding of the material. Given that science is laden with jargon, we feel these results are particularly important in science education, and may be even more relevant when teaching nonmajors, or to students who are learning science in a second language. Our results point toward a need to further explore the effects of jargon, and cognitive load, on student learning and mastering of biological concepts.

Acknowledgements

We are pleased to acknowledge helpful discussions with Trish Schulte and Sarah Gilbert, and the course instructors Sunita Chowrira, Carl Douglas, Marcia Graves, Ehleen Hinze, and Karen Smith; and support of this work by the University of British Columbia through the Carl Wieman Science Education Initiative.

References

[1] Groves, F. H. (1995) Science vocabulary load of selected secondary science textbooks. School Sci. Math. 95, 231–235.

[2] Yager, R. E. (1983) The importance of terminology in teaching K-12 science. J. Res. Sci. Teach. 20, 577–588.

[3] Wandersee, J. H. (1988) The terminology problem in biology education: A reconnaissance. Terminology 50, 97–100.

[4] Wellington, J. and Osborne, J. (2001) Language and Literacy in Science Education. Open University Press, Pennsylvania.

[5] Ryan, J. N. (1985) The language gap: Common words with technical meanings. J. Chem Ed. 62, 1098.

[6] Osborne, J. (2002) Science without literacy: A ship without a sail? Cambridge J. Ed. 32, 203–215.

[7] Yager, R. E., Akcay, H., Choi, A., and Yager, S. O. (2009) Student success in recognizing definitions of eight terms found in fourth grade science textbooks. Electr. J. Sci. Educ. 13, 83–99.

[8] Brown, B. A. and Ryoo, K. (2008) Teaching science as a language: A content-first approach to science teaching. J. Res. Sci. Teach. 45, 529–553.

[9] Lloyd, C. V. and Mitchell, J. N. J. N. (1989) Coping with too many concepts in sciences texts. J. Reading 32, 542–545.

[10] Seifer, K. and Espin, C. (2012) Improving reading of science text for secondary students with learning disabilities: Effects of text reading, vocabulary learning, and combined approaches to instruction learning. Learning and Disability Quarterly 35, 236–247.

[11] Zhang, F. and Lidbury, B. A. (2012) It's all foreign to me: learning through the language of genetics and molecular biology. In Proceedings of The Australian Conference on Science and Mathematics Education (formerly UniServe Science Conference).

[12] Sweller, J. (1988) Cognitive load during problem solving: Effects on learning. Cog. Sci. 12, 257–285.

[13] Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., and Wenderoth, M. P. (2014) Active learning increases student performance in science, engineering, and mathematics. Proc. Natl. Acad. Sci. 111, 8410–8415.

[14] Charmaz, K. (2008) Grounded Theory in Qualitative Psychology: A Practical Guide to Research Methods. Smith, J (Ed). Sage Publications. Thousand Oaks, CA, pp. 81–110.

[15] Kahnemann, D. (1973). Attention and Effort. Prentice-Hall, New Jersey.

[16] Navon, D. and Gropher, D. (1979) On the economy of the human-processing system. Psychol. Rev. 86, 214–255.

[17] Sweller, J. (1994) Cognitive load theory, learning difficulty, and instructional design. Learning and Instruction 4, 295–312.

[18] Clarke, T., Ayres, P., and Sweller, J. (2005) The impact of sequencing and prior knowledge on learning mathematics through spreadsheet applications. Ed. Tech. Res. Dev. 53, 15–24.

[19] Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., and Norman, M. K. (2010) How learning works: Seven research-based principles for smart teaching, Chapter 2. John Wiley & Sons.

[20] Bloom, B. S., Krathwohl, D. R., and Masia, B. B. (1956) Taxonomy of Educational Objectives: The Classification of Educational Goals. New York, McKay.

[21] Ericsson, K. A., Charness, N., Hoffman, R. R., and Feltovich, P. J., Eds. (2006) The Cambridge Handbook of Expertise and Expert Performance, Ericksson. Cambridge University Press, pp. 683–704.