

HOW WE TEACH | *Generalizable Education Research*

Redesigning a course to help students achieve higher-order cognitive thinking skills: from goals and mechanics to student outcomes

Janet Casagrand and Katharine Semsar

Department of Integrative Physiology, University of Colorado Boulder, Boulder, Colorado

Submitted 1 July 2016; accepted in final form 7 February 2017

Casagrand J, Semsar K. Redesigning a course to help students achieve higher-order cognitive thinking skills: from goals and mechanics to student outcomes. *Adv Physiol Educ* 41: 194–202, 2017; doi:10.1152/advan.00102.2016.—Here we describe a 4-yr course reform and its outcomes. The upper-division neurophysiology course gradually transformed from a traditional lecture in 2004 to a more student-centered course in 2008, through the addition of evidence-based active learning practices, such as deliberate problem-solving practice on homework and peer learning structures, both inside and outside of class. Due to the incremental nature of the reforms and absence of pre-reform learning assessments, we needed a way to retrospectively assess the effectiveness of our efforts. To do this, we first looked at performance on 12 conserved exam questions. Students performed significantly higher post-reform on questions requiring lower-level cognitive skills and those requiring higher-level cognitive skills. Furthermore, student performance on conserved questions was higher post-reform in both the top and bottom quartiles of students, although lower-quartile student performance did not improve until after the first exam. To examine student learning more broadly, we also used Bloom's taxonomy to quantify a significant increase in the Bloom's level of exams, with students performing equally well post-reform on exams that had over twice as many questions at higher cognitive skill levels. Finally, we believe that four factors provided critical contributions to the success of the course reform, including: transformation efforts across multiple course components, alignment between formative and evaluative course materials, student buy-in to course instruction, and instructional support. This reform demonstrates both the effectiveness of incorporating student-centered, active learning into our course, and the utility of using Bloom's level as a metric to assess course reform.

Bloom's taxonomy; course reform; alignment; deliberate practice; structured reform; active learning

THE SCIENCE EDUCATION community is encouraging a widespread shift to an evidence-based pedagogy that engages students to become more expert-like in their thinking (1, 34). This requires students to integrate conceptual knowledge and develop discipline-relevant problem-solving skills (1, 36). In addition, there is a call for science educators to provide evidence that instruction results in students achieving these rigorous learning goals (1). Thus science educators at many universities and colleges are being asked to 1) change what they are doing in the classroom, and 2) provide evidence that the changes they are making are translating into higher student learning (e.g., Refs. 1, 34).

The following is a description of a major revision to an upper division, undergraduate neurophysiology course. Over 4 yr, the same instructor (author JC) taught this course and incrementally incorporated evidence-based, active learning approaches. While active learning approaches have a wealth of support that they improve student learning (23, 36, 53), implementation in classrooms is sometimes met with mixed success (4, 39, 43, 49, 50). In addition, the neurophysiology course revisions required both substantial time to develop and refine and increased the workload for students and graduate teaching assistants (TAs). Because of these factors, we sought to justify the added workload and determine whether the implementation of active learning improved student learning and conceptual understanding.

Because the course reforms occurred incrementally over time without the original intention of measuring student learning before and after reform, it was difficult to use standard research techniques to assess changes in student learning such as pre-/post-assessments or student attitude surveys (e.g., Ref. 48). However, we were able to examine student performance on 12 exam questions that were conserved pre- and post-reform. Then, to more broadly examine student learning, we characterized and compared all of the pre-reform and post-reform exam questions using Bloom's taxonomy, a widely accepted tool for delineating the cognitive levels of assessments in terms of lower- to higher-order levels of thinking (e.g., Refs. 6, 14, 19, 54). Bloom's taxonomy of educational objectives for the cognitive domain (7) identifies six levels of understanding: knowledge/remember, comprehension/understand, application/apply, analysis/analyze, synthesis/create, and evaluation/evaluate (2, 7), with the first two levels generally considered to represent lower levels of understanding [lower-order cognitive skills (LOCS)] and the other four, higher-order levels involving critical thinking [higher-order cognitive skills (HOCS)] (e.g., Refs. 6, 14, 19, 54). Thus Bloom's taxonomy is a useful tool for evaluating and quantifying the cognitive skill level of the exams before and after the course transformation and an appropriate tool to indirectly address whether the course reforms were effective in improving student learning.

To also examine reasons that could explain why reforms appeared to be improving student learning, we first characterized the Bloom's level of questions on the new, formative assessments to determine their degree of alignment with the revised exams. We also surveyed students in the post-reform semester to gauge how helpful students thought the course reforms were for their learning.

Address for reprint requests and other correspondence: J. Casagrand, Dept. of Integrative Physiology, University of Colorado, 354 UCB, Boulder, CO 80309-0354 (e-mail: Janet.Casagrand@colorado.edu).

METHODS

This study was conducted under the Institutional Review Board protocol no. 0108.9 (exempt status).

Course description and student demographics. The neurophysiology course described here is an upper division (junior and senior level), one-semester course for Integrative Physiology majors at the University of Colorado who are generally interested in the allied health professions. In both semesters considered in the study (Fall 2004 = pre-reform; Fall 2008 = post-reform), the course consisted of three, 50-min lectures per week and three exams composed of multiple choice and short answer questions. The course was taught by the same instructor (JC), and the basic course content was the same both semesters and included an overview of nerve cells and behavior, neural signaling and its basis (electrical properties, driving forces, current flow, passive membrane properties, sensory transduction, synaptic transmission and plasticity), and the neural basis for movements (motor units and muscle force, spinal reflex circuits, locomotion and rhythmic behaviors, and voluntary movements). Course prerequisites also did not change and included 1 yr of general biology, 1 yr of chemistry, one semester of statistics, and one semester of physiology. In addition to lecture, there was a required 2-hr laboratory/recitation each week with one-half of the time devoted to recitation and one-half to a laboratory activity. While the laboratory activities did not differ between the semesters, there was a change to how the recitation time was spent. In both pre- and post-reform semesters, time was devoted to addressing student questions about lecture material. However, in the post-reform semester, the recitation also included time to discuss homework questions submitted the previous week.

Typical enrollment in the course varies from 80 to 110 students per semester. In the two semesters considered in the study, there were 82 students in the pre-reform semester and 97 in the post-reform semester. The student populations in both semesters were similar (Table 1) with no significant differences in class standing (i.e., percentage of juniors vs. seniors), cumulative college grade point average (GPA) of students when entering the course, or GPA in major. There was also no significant difference in cumulative GPA between the two semesters for the top quartile or the bottom quartile. We did observe a difference in the ratio of male to female students in the two semesters, with fewer female students in the post-reform semester. However, for neither semester was there a difference between the male and female students in cumulative GPA, GPA in major, or exam performance.

Course revisions. Before 2005, the neurophysiology course was taught in a teacher-centered, traditional lecture format. Although the instructor's goal for students was to have them be able to apply

neurophysiology concepts in new situations and develop more expert-like thinking, not simply memorize facts, it was difficult to challenge students' understanding of the material on exams, despite spending considerable lecture time on the concepts. In many cases, students could only reproduce information from lecture. For example, students could reproduce a drawing of the phases of an action potential but were not able to predict what effect a potassium or sodium channel blocker would have on the phases of an action potential.

To try to improve students' abilities to apply neurophysiology concepts, the course was iteratively revised over 4 yr (starting in 2005) by gradually adding evidence-based course reforms (e.g., Refs. 15, 22, 23, 53). While the basic course content was consistent between semesters, the post-reform course placed less emphasis on memorization and more on understanding the content at a more conceptual level. With this shift in philosophy, several different types of course reforms were implemented to provide students structured support with which to meet these goals (Table 2). Each reform, including homework assignments, a homework help room, in-class clicker questions, and explicit learning goals, was designed to serve a specific purpose.

First, in 2005, homework assignments were added to give students both feedback on their mastery of concepts and deliberate practice working with concepts with which they often struggled on exams. This was the first reform because students were already requesting practice with course content before exams. The initial assignments were fairly short (requiring ~15–20 min to complete) and were designed to give students practice working with HOCS questions, including solving problems and articulating their reasoning (17). Assignments also required students to work with LOCS, such as retrieval of facts and vocabulary, which were necessary to complete HOCS questions. Based on positive student responses, these assignments were expanded over the next 3 yr with questions being revised and added. By 2008, homework assignments were due weekly (with the exception of weeks when there was an exam), took a couple of hours to complete, and comprised 15% of a student's final grade. Graded homework assignments were reviewed in the following week's recitation (a change to how recitation time was allocated).

While students were always encouraged to work together on homework assignments, in 2008 an infrastructure for peer collaboration and discussion was added by creating an informal, optional homework help room based on principles of peer learning (13, 29, 33). The homework help room was a designated time and place for peers to help one another, with a graduate TA available or the instructor present to answer questions as needed. The homework help room was available for several hours the day before the homework assignment was due. Most students completed the assignment before coming in and then worked in small groups to discuss their answers and reasoning, particularly for questions they were unsure about. When students got stuck or disagreed with one another, the instructor or graduate TA was available to help socratically guide student discussions.

In 2006, the instructor started using a personal response system (a.k.a., "clickers") to improve student engagement in lecture (15, 33). Clicker questions were presented as multiple-choice questions on a PowerPoint slide. Initially, each lecture averaged 1.5 clicker questions ($n = 55$ per semester) with a mix of LOCS and HOCS questions. Developing both the homework and clicker questions was an iterative process. Initially homework and clicker questions were sourced from student questions during lecture and office hours, as well as student difficulties and misconceptions on homework assignments and exams. As students responded favorably and reported finding the questions helpful for their learning, the number of clicker questions increased each semester. By Fall 2008, each lecture included 2–6 clicker questions (average of 4 questions per lecture; 177 questions total per semester). Clicker question formats were mostly HOCS and had evolved into what is now recognized as best practice, with students having time to discuss the question with neighbors, report out reasoning, followed by instructor explanations (9, 47). Students received two participation points for answering the question and a bonus point for

Table 1. Comparison of student demographics for the pre-reform and post-reform semesters

	Pre-reform ($n = 82$)	Post-reform ($n = 97$)	Statistical Difference
Class standing, %			$\chi^2 = 10.15; P = 0.07$
Junior	22	18	
Senior	72	81	
Cumulative college GPA			
Overall	3.0 ± 0.51	3.1 ± 0.53	$P = 0.22$
Top quartile	3.7 ± 0.18	3.7 ± 0.15	$P = 0.12$
Bottom quartile	2.4 ± 0.27	2.4 ± 0.17	$P = 0.78$
Sex ratio (male/female), %	32:68	50:50	$\chi^2 = 6.8, P = 0.009^*$

Cumulative college GPA is given as means \pm SD; n , no. of subjects. GPA data were not available for a few students (Pre, $n = 4$; Post, $n = 2$). Sex was not known for 4 students in the post-reform semester. *While there was a difference in the ratio of male and female students, between male and female students there was neither a difference in cumulative GPA (t -test, Pre: $P = 0.14$; Post: $P = 0.31$), nor exam performance (t -test, Pre: $P = 0.71$; Post: $P = 0.74$).

Table 2. Course reforms were incrementally added over 4 yr

	F04 (Pre-reform)	F05	F06	F07	F08 (Post-reform)
Lecture	3 × 50 min/wk	→			
Laboratory/recitation	110 min/wk	→			
Homework		15–20 min	(average time spent/wk)		2–6 h
Clickers			1.5	(average no. of questions/lecture)	4
Homework help room					3 h
Explicit learning goals					Provided (6 course goals; 50 topic goals)

In Fall 2004 (F04; pre-reform), the course was taught in a traditional lecture format. Homework assignments were introduced in Fall 2005 (F05), clicker questions in Fall 2006 (F06), and a homework help room and explicit learning goals in Fall 2008 (F08) (post-reform).

answering the question correctly, accounting for 5% of the final grade (see Ref. 15).

In Fall 2008, learning goals were also introduced into the course to make expectations explicit to both the instructor and the students. These learning goals included 3 concept-related course goals, 3 skills-related course goals, and ~50 content-specific learning objectives (Table 3). The instructor provided all course goals and learning objectives to students, encouraging students to use them as a guide to what was expected of them on exams.

Exam revisions. As a result of these incremental changes, the course exams also gradually changed from 2004 to 2008. By Fall 2008, only 14% of questions ($n = 12$) remained the same as on the Fall 2004 exams. The reason for these changes was not because of a difference in course content (i.e., topics covered), which was highly conserved, but rather to 1) reflect the instructor's intention to have students demonstrate a more conceptual understanding of the material; 2) align exams with the homework and clicker questions, and

explicit learning goals that were based not only on content, but also on Bloom's level; and 3) maintain a similar exam average. With regards to the last point, after each exam, the instructor used statistical analyses to determine which questions were discriminating well and modified the exams to eliminate or revise questions that were not discriminating well, including eliminating questions that a high percentage of students answered correctly. As intended, exam means on all but the first exam were similar between pre- and post-reform semesters (Table 4). Finally, for all semesters across the 4 yr, exams were not returned to students. Students were required to come into the instructor's office to review their exam.

Assessing effectiveness of course reforms. To determine effects of the aforementioned course revisions on student learning of neurophysiology, we performed the following analyses. First, we compared student performance on LOCS and HOCS exam questions that were conserved before and after course reforms. In addition, to determine whether the reforms differentially affected student performance for higher- or lower-performing students, we compared performance on these conserved questions before and after reform among the top and bottom quartiles of scores separately. Top and bottom quartiles were identified separately for each exam.

Second, because the sample size of conserved questions was small [although comparable to many concept inventories (see DISCUSSION; Refs. 27, 38, 46)], we also used Bloom's taxonomy to assess changes in cognitive levels of exams to get a more complete picture. To decrease bias in the Blooming process, we recruited three independent raters who were sufficiently familiar with the subject matter: two former graduate students currently employed as science teaching fellows (STFs), and one current graduate student. Although raters knew they were categorizing neurophysiology course questions, they were blind to the semester and purpose of the study. Raters were trained in Bloom's taxonomy and use of a Bloom's Dichotomous Key (BDK) developed by Semsar and Casagrand (41a). Before raters categorized questions, we first had the course instructor answer *question 1* (Q1) on the BDK (i.e., could students memorize the answer to this specific question?) for all exam questions from both semesters. The instructor based this decision on whether the answer to the question had been taught and could simply be memorized. For example, in lecture, Ohm's law is presented (i.e., $V = I \times R$). If the question were "What is the equation for Ohm's law?" or "What variables are included in the Ohm's law equation?", then students could memorize the answer to the question. However, if the question was "Calculate how much current flows under the following conditions...", students would need to use their knowledge of Ohm's law to answer the question, and the answer to that specific question could not be memorized. Once the instructor had answered Q1 of the BDK for all the exam questions, we then had raters use the BDK to categorize the exam questions into one of the six Bloom's categories: 1) remember, 2) comprehend, 3) apply, 4) analyze, 5) synthesize/create, and 6) evaluate. Interrater reliability was high, with a Krippendorff's α of 0.79 (24).

Table 3. Examples of course goals and specific learning objectives in the neurophysiology course

Course Goals – Content Orientated
1) Predict and explain how the flow of ions across the nerve cell membrane can produce and influence the signals used in the nervous system to communicate information (both within and between neurons).
2) Predict and explain how information in the nervous system is converted from one type of signal/information to another, and how the properties of neurons can influence this process.
3) Predict and explain how the properties of individual neurons, and the types and patterns of connections between neurons, can influence activity in the nervous system, can influence behaviors (as demonstrated through basic types of movement), and can be adjusted, adapted, or altered to suit the changing needs of an organism.
Course Goals – Skill Orientated
1) Be able to hypothesize and state the connections between concepts in nervous system function, instead of simply memorizing facts, as a way for you as a student to better understand nervous system function.
2) Be able to interpret and evaluate scientific data collected with techniques commonly used in neurophysiology to better understand concepts of nervous system function.
3) Improve problem solving skills to help in understanding concepts and predicting aspects of nervous system function.
Examples of Specific Learning Objectives
1) Predict the neural signal you would expect to see in the four functional regions of the three types of neurons for a novel scenario.
2) Calculate the net driving force or the equilibrium potential for an ion, or the resting membrane potential of a cell.
3) Given a change in a parameter that influences net driving force, equilibrium potential, or resting membrane potential, predict how this would influence net driving force, equilibrium potential, resting membrane potential, or current across the membrane.

Table 4. Exam scores before and after course reform

	Pre-reform	Post-reform	Statistical Difference
Exam 1	73.7 ± 11.6	66.6 ± 16.3	$P < 0.01^*$
Exam 2	75.3 ± 13.5	74.1 ± 13.2	$P = 0.56$
Exam 3	70.5 ± 12.0	70.1 ± 14.2	$P = 0.81$

Scores are given as means ± SD. Exam 3 was not cumulative. *Significantly different from pre-reform semester at $P < 0.05$; t -test.

For statistical analysis of Bloom's levels, we first assigned a single Bloom's value (1–6) to each question. To do this, we applied the decision rules published by Zheng et al. (54). Briefly, if at least two of the raters agreed on the same Bloom's level, that ranking was assigned to the question. If two raters did not agree, we did one of two things. When ratings were sequential (e.g., 2–3–4), we assigned the middle value. When ratings were nonsequential (e.g., 1–2–4), we assigned the arithmetic average. We then assigned noninteger values a category, depending on whether they were within one-half point of the category (rounding up at 0.5). These values were then averaged across all exams in a semester to arrive at three different metrics for comparison: an unweighted Bloom's average, a simple weighted Bloom's average (weighting each question by the point value assigned to that question, see Ref. 54), and the weighted Bloom's index (weighting each question by their relative contribution to a student's total possible score, see Ref. 19).

To judge the degree of alignment of course exams and course materials, we also quantified the Bloom's taxonomic level of all course materials in the post-reform semester, including homework questions ($n = 217$) and clicker questions ($n = 177$). As the number of materials was high, only a single rater was used to rate the homework and clicker questions.

Finally, to examine student perceptions of the course changes and how useful students found various course activities, we surveyed students in 2008 via an online, anonymous, end-of-semester survey. The survey was administered and analyzed by an independent third-party. Students were asked about their attitudes toward various components of the course, such as how helpful for their learning they found the components (e.g., homework assignments, clicker questions, help room) and how much they enjoyed them.

Data analysis. Data were analyzed statistically using either a t -test or χ^2 test, as appropriate, as noted in the RESULTS. A level of significance of 0.05 was used. Normalized learning gains were calculated using the method found in Fagan et al. (18), and effect sizes were calculated using the method of Cohen (10).

RESULTS

With the exception of the lower-performing students on the first exam, overall students were performing as well on post-reform exams as they had before the course improvements (Tables 4 and 5).

Conserved exam questions. On the 12 conserved exam questions, student performance was significantly higher in the post-reform semester ($P = 0.002$; t -test), changing from 59% (± 18.8 SD) pre-reform to 76% (± 19.1 SD) post-reform. This represents a normalized learning gain of 42% and an effect size of 0.9.

After Blooming questions, we further separated these conserved questions into LOCS (recall and comprehend; $n = 5$) and HOCS (apply, analyze, evaluate, synthesis/create; $n = 7$) and compared the percentage of students in the two semesters who correctly answered each type of question (Fig. 1). We observed that, for both LOCS and HOCS questions, students in the post-reform course performed significantly better, includ-

Table 5. Top and bottom quartile student performance in the pre-reform and post-reform semesters for the three exams

	Pre-reform, %	Post-reform, %
Exam 1		
Top	86.5 ± 3.8	85.3 ± 5.8
Bottom	57.7 ± 6.3	44.9 ± 7.4*
Exam 2		
Top	89.6 ± 4.3	89.2 ± 4.2
Bottom	56.0 ± 8.7	56.6 ± 7.9
Exam 3		
Top	84.3 ± 5.6	85.3 ± 3.4
Bottom	55.0 ± 5.9	51.4 ± 10.9

Scores are given as means ± SD. *Significantly different from pre-reform semester at $P < 0.05$; t -test.

ing 21% better on LOCS questions (t -test; $P = 0.03$; normalized learning gain = 60%, effect size 1.0) and 13% better on HOCS questions (t -test; $P = 0.02$; normalized learning gain = 44%, effect size 0.7).

To determine whether the course reforms were differentially helping higher-performing or lower-performing students, we compared the pre- and post-reform performance of the top and bottom quartiles of the class on the six conserved questions on the first exam and the six conserved questions combined from the second and third exams. Among higher-performing students, student performance in the post-reform semester was 10% higher on conserved questions on both the first exam (t -test: $P = 0.02$, normalized learning gain = 50%, effect size 1.2) and 13% higher on the second and third exams (t -test; $P = 0.04$; normalized learning gain = 62%, effect size 0.6), representing an effect size of 1.2 and 0.6 (Fig. 2). Among lower-performing students, there was no difference in performance on the first exam (t -test; $P = 0.44$; normalized learning gain = 1%; effect size = 0.05). However, student performance in the post-reform semester was 34% higher on the second and third exams (t -test; $P = 0.009$; normalized learning gain = 54%; effect size 1.3). Not only did the lower-performing students in the post-reform semester have higher scores on

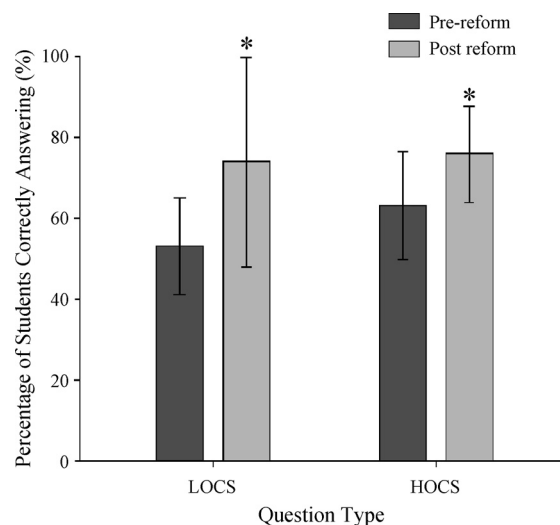


Fig. 1. Student performance on conserved lower- (LOCS, $n = 5$) and higher-order (HOCS, $n = 7$) exam questions in the pre-reform and post-reform semesters. Scores are given as means ± SD. *Significant difference from the pre-reform semester.

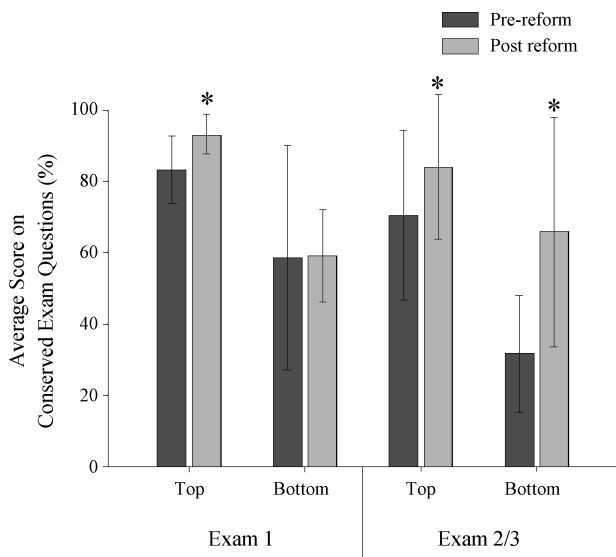


Fig. 2. Performance of higher- and lower-performing students on the conserved exam questions in the pre-reform and post-reform semesters for the first exam ($n = 6$ questions) vs. the second and third exams ($n = 6$ questions). Scores are given as means \pm SD. *Significant difference ($P < 0.05$) from the pre-reform semester.

conserved questions on the latter two exams, the number of students answering these questions correctly doubled from 32% to 66%.

Blooming exams and course materials. As shown in Fig. 3, post-reform exams had significantly more questions at the higher Bloom's levels ($\chi^2; P < 0.001; n = 83, 84$) with the average, unweighted Bloom's level of 1.8 pre-reform and 3.3 post-reform (Table 6), effect size = 1.2. The weighted Bloom's level (i.e., each question's Bloom's category weighted to reflect its point value on the exam) mirrored this change (Table 6). Overall, the number of questions falling into the HOCS categories more than doubled, increasing from 24% to 67%. This was largely due to a dramatic decrease in the number of *remember* questions and a large increase in the number of *analyze* questions, as well as the introduction of

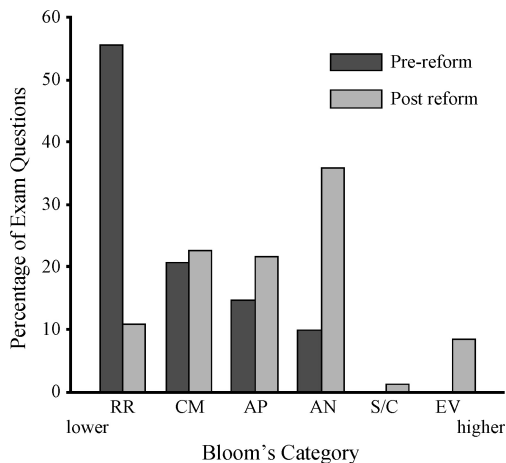


Fig. 3. Comparison of the percentage of exam questions in each of the Bloom's categories (RR, remember; CM, comprehend; AP, apply; AN, analyze; S/C, synthesize/create; EV, evaluate). Post-reform exams had significantly more questions at the higher Bloom's levels.

Table 6. Average Bloom's level for pre- and post-reform exams

	Pre-reform	Post-reform
Unweighted*	1.8	3.3
Weighted†	1.9	3.5
Weighted Bloom's Index‡	31.4	57.4

*Average of each question's Bloom's category (54). †Average of each question's Bloom's category weighted to reflect its point value on the exam (54). ‡Weighting each question by its relative contribution to a student's total possible score (19).

several *create* and *evaluate* questions (Fig. 3). These changes were present on all three exams, as the weighted Bloom's levels of each exam increased post-reform (*exam 1*: 2.1 to 3.2, *exam 2*: 2.1 to 3.7, *exam 3*: 1.6 to 3.3).

The Bloom's level of post-reform course materials was compared with the Bloom's level of exam questions to demonstrate the degree of alignment between the formative and evaluative course materials (Table 7). The in-class clicker questions focused on having students demonstrate comprehension of concepts (42%), while homework questions focused on higher order skills of application and analysis (57%) with some create/synthesis and evaluation (15%).

Student perceptions. The post-reform survey in Fall 2008 had a response rate of 90%. In general, the survey indicated that students enjoyed the course activities and felt they helped their understanding and exam preparation. For example, students reported using the learning goals as a study tool to help them focus on the most important material, to organize their notes, or for self-quizzing. Seventy percent of respondents reported using the help room, and all recommended having one in the future. Ninety-one percent of respondents enjoyed the in-class clicker questions, and 90% found them helpful for their understanding. Fifty percent enjoyed the homework assignments, and 22% were neutral. One hundred percent of respondents reported the homework helped their understanding of the course material, including 98% reporting the homework helped with exam preparation and 80% reporting that they felt they did not have to study/cram as much for exams. Students also reported spending an average of 2–6 h per week outside of lecture on the course.

DISCUSSION

We believe that many of the lessons learned here will be helpful to instructors who are implementing reform but are still new to scientific pedagogy. In addition, we hope this reform description will be useful to those studying the implementation of scientific teaching by describing factors that were important

Table 7. Distribution of question types on the practice assignments and post-reform exams

	n	Bloom's Category, %					
		RR	CM	AP	AN	SC	EV
Clicker	177	18	42	29	11		
Homework	217	2	26	40	17	5	10
Post-reform exam	84	11	23	21	36	1	8

n, No. of subjects. Bloom's categories: RR, remember; CM, comprehend; AP, apply; AN, analyze; SC, synthesize/create; EV, evaluate.

to the successful transformation presented here. First, we demonstrate significantly improved student learning in a transformed neurophysiology course using two methods: analyzing student performance on conserved exam questions, and comparing Bloom's levels of exams. Second, we discuss the key aspects of reform that were instrumental to its success.

Reform was worth the effort. Our primary goal with this work was to determine whether the course revisions were broadly leading to deeper student learning, making them worth the increase in time investment for students, graduate TAs, and the instructor. The effectiveness of using active learning to enhance student learning was supported by student performance on the 12 conserved exam questions. While 12 questions is a small sample, many validated concept inventories designed to assess the effectiveness of changes in instruction are in the range of 11–30 (e.g., Refs. 27, 38, 46). Therefore, while our conserved questions represent a small sample set, we believe they still provide meaningful results. On these conserved questions, student overall performance was significantly higher after the course revisions, with effect sizes greater than the average effect size (0.47) reported by Freeman et al. (20) and in general a very large effect size for education studies (10, 39).

While the reforms were designed to improve higher order learning, results from the conserved questions also suggested students made learning gains in both LOCS and HOCS. We believe that the improvement on LOCS questions may have been due to careful inclusion of LOCS elements within HOCS questions. This may have provided students with context and necessity to acquire the LOCS and prioritize acquisition of this knowledge (35, 37). For example, consider the HOCS question, "What are the directions of the chemical, electrical, and net driving forces acting on K^+ when the membrane potential is +55 mV?" Answering this requires students to first know what these forces are and whether they are, or are not, influenced by changing the membrane potential (LOCS) and then apply that understanding to this specific situation (HOCS). In addition, as students were getting more structured practice with the concepts and required problem-solving, they were likely working more with the required factual knowledge and thus better able to retrieve such information (35, 37).

Additionally, results from the conserved questions suggested course revisions were helping both lower- and higher-performing students, with the higher- and lower-performing students having comparable learning gains on the latter exams post-reform (Fig. 2). However, while the top-performing students responded by the first exam, the low-performing students took longer to show improvement. Active learning has previously been shown to help both higher-performing (5) and lower-performing students (5, 19, 21). Here we demonstrate that post-reform student performance is significantly higher on the conserved questions of latter exams for students in both the top and bottom quartiles of the course (Fig. 2). This leads to a striking improvement for lower-performing students, as pre-reform they only had 32% of questions correct, while post-reform this doubled to 66%. It may be that the structure of the course supported the learning of the lower-performing students, as increased structure has been shown to decrease failure rates (19). Interestingly, however, while the higher-performing students were already performing better on the first exam, the lower-performing students did not show improvement until the

latter exams, which were of equivalent or greater cognitive level. Thus, while a structured learning environment helps lower-performing students, it may take them more time to adapt their study style and embrace the learning opportunities the reforms provided.

In addition to the results of the conserved exam questions, the overall Bloom's cognitive skill level of all post-reform exam questions more than doubled compared with pre-reform questions (Fig. 3), again representing a large effect size (1.2). As students were performing equally well on these higher-level exams, this is indirect evidence that the course revisions substantially increased student learning in the course. Given the large success of the course reform in allowing students to achieve the ability to perform well on HOCS exams, we will further look into why this course reform was so successful.

Keys to success of course reform. Despite a great deal of research indicating that active learning techniques help students learn more and at deeper conceptual levels (20), several key barriers to reform have been identified, including student push-back (41, 43), lack of supportive pedagogical infrastructure (8, 26), and lack of professional reward and promotion structures (16, 25, 52). If instructors and/or students cannot see the benefits associated with the extra time and effort active learning requires, too often instructors will abandon, rather than refine, their initial attempts at course reform (43, 50, 51). While the course reform chronicled here includes well-documented active learning techniques, its success and large student learning gains in HOCS warrant reflection on why the changes were successful and why they continue to be in place today.

We believe the success of the course reform is largely based on its holistic inclusion of multiple evidence-based teaching strategies. While course revisions were introduced incrementally and continually improved upon, by the end of the 4th yr of course reform many of the fundamental elements of evidence-based pedagogy were in place, such as the loop of guidance/practice/feedback, course alignment, student buy-in, and support infrastructures for both students and the instructor. First, students received more guidance, practice, and frequent and timely feedback, three critical components of successful active learning strategies (12, 23). Through the introduction of concept-based homework, a homework help room, in-class clicker questions, and explicit learning goals based not only on content but also Bloom's level, students received more guidance through instructor-modeled problem-solving strategies in both lecture and the homework help room (23, 42). Students also received more practice working at the higher cognitive levels, both in lecture through the use of clicker questions and outside of class through challenging homework activities (17). Furthermore, the ability to discuss clicker questions in lecture and work together on problem-solving techniques in the homework help room provided a substantial infrastructure for frequent and timely feedback from the instructor and self-assessment for students to gauge their level of understanding (e.g., Refs. 47, 53). Finally, these opportunities also provided extensive infrastructure for peer learning, a strong influence on HOCS (13, 32).

Second, there was strong alignment of the course goals, formative assessments, and evaluative assessments in the course. Because assessments inform students of what they need to know and do, new learning opportunities need to be aligned well with assessments to be successful (see Refs. 3, 11, 40). As

the course progressed, explicit learning goals were developed that drove the creation of exam questions and accompanying course materials (examples, Table 3; Refs. 44, 45). These learning goals included not only what students needed to know (content knowledge; e.g., equations for Ohm's law, equilibrium or resting membrane potentials), but also what students needed to be able to do (skill level; e.g., calculate a value, predict a change, etc.). Assuring that the instructional activities were matched to the level at which the students were expected to perform on exams was an important part of the success students achieved in the reformed course.

Third, we had significant student buy-in for the course revisions. Although students may not be best at recognizing what helps them learn or in assessing their true level of understanding (28, 30, 31), student attitudes are an indicator of student buy-in to teaching strategies. In turn, student buy-in and willingness to make use of the learning opportunities provided by an active learning classroom is critical to its success (35). On our end-of-term survey following the course reform, students overwhelmingly responded positively to the course. Over 90% of responding students both enjoyed and found in-class clicker questions helpful for their understanding, and all found homework helpful for their learning. Perhaps even more telling was the high use of optional resources, such as the homework help room. Seventy percent of responding students reported using the homework help room, much higher than the estimated 10–20% of students who previously came to instructor/TA office hours. In addition, of the 70% of students using the help room, nearly all (93%) reported working with other students, indicating that not only were opportunities for peer learning provided, but students were utilizing these resources at a high level. Finally, students also reported finding the learning goals helpful as a study tool to help them focus on the most important material, to organize their notes, or for self-quizzing. In summary, not only were students performing

at much higher cognitive levels, but they were accepting, enjoying, and using the tools that help them achieve those goals.

A final piece of the course reform's success was the departmental/university level support for course reform, the lack of which is a strong barrier to reform. While course reforms began before any departmental support was available, in Fall 2006 the Integrative Physiology department became involved with the Science Education Initiative (SEI) on the University of Colorado Boulder campus and hired three STFs (52). The program provided several levels of support to the instructor, including SEI-sponsored workshops on evidence-based pedagogy that the instructor attended and the help of the STFs for course development. For example, the STFs helped the instructor to refine course goals and strategies and helped the instructor to develop explicit learning goals for the course, intentionally focusing on the alignment of course materials. The STFs were also able to help introduce the instructor to new techniques and approaches, provide emotional support when difficulties arose, and help administer and analyze student surveys. Additionally, with the encouragement and support of the STFs, the instructor applied for and received a University of Colorado Boulder President's Teaching and Learning Collaborative award in 2008 to develop a way of measuring the effectiveness of the course reform efforts retrospectively. This award provided modest resources that allowed the instructor to collect data to inform instructional changes and sustain the reform effort.

Additional changes since 2008. Changes have continued to be made to the course based on the results documented above. Several of the more major changes are noted here. First, because lower-performing students take time to adapt to the changes, the exam structure has been altered for the course. Initially the course had three, equally weighted exams. Since 2008, a fourth exam has been added, and a student's lowest

Table 8. *Examples of alignment between learning goals, practice (clicker, homework) questions, and exam questions*

Example	Learning Goal	Practice (Clicker, Homework) Question	Exam Question
1	Given a change in a parameter that influences net driving force, <i>equilibrium potential</i> , or <i>resting membrane potential</i> , predict how this influences net driving force, <i>equilibrium potential</i> , <i>resting membrane potential</i> , or current across the membrane.	If the external concentration of sodium is doubled, will the <i>resting membrane potential</i> change? If the sodium permeability is transiently doubled, will the <i>resting membrane potential</i> change?	A drug is applied to a neuron that temporarily blocks a substantial portion (~65%) of the resting K ⁺ ion channels in the neuron's membrane preventing any ions from moving through these channels. <ul style="list-style-type: none"> • What effect, if any would this drug have on the <i>equilibrium potential</i> of K⁺ in this neuron? • What effect, if any, would this drug have on the <i>resting membrane potential</i> of the neuron?
2	Predict the type of ion channel in a given functional region of a neuron for a novel scenario.	What type of gated ion channel would you most likely expect to find at the input region of the motor neuron (arrow) in the figure above?	Considering the 4 types of gating stimuli, which type of gated ion channel would you expect to be characteristic of the input region of this sensory afferent?
3	Calculate time and length constants, input resistance and input capacitance from graphical data (for ex., I–V data) or numerical values	Calculate the input resistance for this neuron, based on these I–V data. Calculate the conductance(s) for this ion channel (based on the I–V data shown).	Calculate the conductance(s) for this ion channel (based on the I–V data shown).
4	Differentiate between electrical, chemical, and net driving forces, and, given a set of data, predict the direction or relative magnitude of the current flow (net or ionic) due to these forces and its effect on the membrane potential.	If the membrane potential is –60 mV and the equilibrium potential for chloride is –80 mV in a typical cell, what are the directions of the electrical, chemical, and net driving forces acting on chloride?	What are the directions of the chemical, electrical, and net driving forces acting on K ⁺ in a typical neuron when the membrane potential is 55 mV?

exam score is weighted one-half that of the other three exams, allowing students to perform poorly on one exam without it significantly affecting their overall performance in the course. The instructor feels this is an important course revision that rewards students who put forth the effort to modify their study practices into more effective learning experiences. Second, to further support students who take longer to adapt to the course format and help them better understand course expectations, a two-part, in-class activity was added that illustrates learning goal/exam alignment. Before the first exam, students are asked to match a subset of questions to a learning goal. Then after the first exam, the instructor matches examples of the exam questions with learning goals, clicker questions, and homework questions to demonstrate alignment and relevance of course learning activities (Table 8). Third, to help manage course resources by decreasing TA grading load, some open-response homework questions were converted to multiple choice questions that can be administered and graded online. Finally, as the homework help room was so popular with students, and students were utilizing it well by working with each other, the instructor expanded the number of help room hours to better accommodate student schedules. It is also worth noting that, perhaps as a side-effect of students utilizing this resource, students appearing in the instructor's office hours are more focused on how to improve their study and exam-taking strategies, and rarely have questions about course material. This has led to many personalized discussions about study skills to help improve the efficiency of students' study habits.

Conclusion. While the improvement of student learning with the introduction of active learning techniques is not novel, it is unfortunately also not a guaranteed result following implementation in the classroom. The active learning environment in this neurophysiology course not only significantly improved higher-order learning in class overall but also promoted increased learning on lower cognitive skills. Furthermore, the active learning classroom helped both higher- and lower-performing students. While students in the bottom quartile of the class did not show improvement initially, they performed much better on the latter exams, indicating that giving students opportunities to adapt to new learning environments is important. The success of these reforms stemmed from a holistic inclusion of evidence-based teaching strategies, such as active learning, course alignment, student buy-in, and support structures for students and the instructor. Finally, in addition to showing a course reform that helped students move from LOCS to HOCS, our approach to quantifying course reform can be a model for instructors to retroactively assess how their course changes have impacted student learning. For instructors who have been making changes without first putting scientific teaching assessment methods in place, Blooming course materials is a helpful approach to measure effectiveness of course reforms retrospectively.

ACKNOWLEDGMENTS

We thank Françoise Benay-Bentley, Dr. Teresa Foley, Dr. Jeffrey Gould, Dr. Dale Mood, Dr. Carl Wieman, and the University of British Columbia CWSEI reading group for assistance with this project.

GRANTS

Financial support was provided by the President's Teaching and Learning Collaborative of the University of Colorado, and the University of Colorado Science Education Initiative.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

J.C. conceived and designed research; J.C. performed experiments; J.C. and K.S. analyzed data; J.C. and K.S. interpreted results of experiments; J.C. prepared figures; J.C. drafted manuscript; J.C. and K.S. edited and revised manuscript; J.C. and K.S. approved final version of manuscript.

REFERENCES

1. **American Association for the Advancement of Science.** *Vision and Change in Biology Undergraduate Education: A Call to Action—Final report* (Online). www.visionandchange.org [18 December 2016].
2. **Anderson LW, Krathwohl DR, Bloom BS (Editors).** *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. San Francisco, CA: Addison Wesley Longman, 2001.
3. **Anderson LW.** Curricular alignment: a re-examination. *Theory Pract* 41: 255–260, 2002. doi:10.1207/s15430421tip4104_9.
4. **Andrews TM, Leonard MJ, Colgrove CA, Kalinowski ST.** Active learning not associated with student learning in a random sample of college biology courses. *CBE Life Sci Educ* 10: 394–405, 2011. doi:10.1187/cbe.11-07-0061.
5. **Beichner RJ, Saul JM, Abbott DS, Morse JJ, Deardorff DL, Allain RJ, Bonham SW, Dancy MH, Risley JS.** Per Central. The Student-Centered Activities for Large-Enrollment Undergraduate Programs (SCALE-UP) Project. www.compadre.org/PER/per_reviews [8 June 2016].
6. **Bissell AN, Lemons PP.** A new method for assessing critical thinking in the classroom. *Bioscience* 56: 66–72, 2006. doi:10.1641/0006-3568(2006)056[0066:ANMFAC]2.0.CO;2.
7. **Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR.** Taxonomy of educational objectives: the classification of educational goals. In: *Handbook I: Cognitive Domain*. New York: David McKay, 1956.
8. **Brownell SE, Tanner KD.** Barriers to faculty pedagogical change: lack of training, time, incentives, and . . . tensions with professional identity? *CBE Life Sci Educ* 11: 339–346, 2012. doi:10.1187/cbe.12-09-0163.
9. **Caldwell JE.** Clickers in the large classroom: current research and best-practice tips. *CBE Life Sci Educ* 6: 9–20, 2007. doi:10.1187/cbe.06-12-0205.
10. **Cohen J.** *Statistical Power Analysis for the Behavioral Sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum, 1988.
11. **Cohen SA.** Instructional alignment: searching for a magic bullet. *Educ Res* 16: 16–20, 1987. doi:10.3102/0013189X016008016.
12. **Collins A, Brown JS, Newman SE.** Cognitive apprenticeship: teaching the crafts of reading, writing, and mathematics. In: *Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser*, edited by Resnick LB. Hillsdale, NJ: Lawrence Erlbaum, 1989.
13. **Crouch CH, Mazur E.** Peer instruction: ten years of experience and results. *Am J Phys* 69: 970–977, 2001. doi:10.1119/1.1374249.
14. **Crowe A, Dirks C, Wenderoth MP.** Biology in bloom: implementing Bloom's Taxonomy to enhance student learning in biology. *CBE Life Sci Educ* 7: 368–381, 2008. doi:10.1187/cbe.08-05-0024.
15. **Duncan D.** *Clickers in the Classroom: How to Enhance Science Teaching using Classroom Response Systems*. San Francisco, CA: Pearson Education, 2005.
16. **Ebert-May D, Derting TL, Hodder J, Momsen JL, Long TM, Jardeleza SE.** What we say is not what we do: effective evaluation of faculty professional development programs. *Bioscience* 61: 550–558, 2011. doi:10.1525/bio.2011.61.7.9.
17. **Ericsson KA, Krampe RT, Tesch-Romer C.** The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 100: 363–406, 1993. doi:10.1037/0033-295X.100.3.363.
18. **Fagen A, Crouch CH, Mazur E.** Peer instruction: results from a range of classrooms. *Phys Teach* 40: 206–209, 2002. doi:10.1119/1.1474140.
19. **Freeman S, Haak D, Wenderoth MP.** Increased course structure improves performance in introductory biology. *CBE Life Sci Educ* 10: 175–186, 2011. doi:10.1187/cbe.10-08-0105.
20. **Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP.** Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci USA* 111: 8410–8415, 2014. doi:10.1073/pnas.1319030111.

21. Haak DC, Hille Ris Lambers J, Pitre E, Freeman S. Increased structure and active learning reduce the achievement gap in introductory biology. *Science* 332: 1213–1216, 2011. doi:10.1126/science.1204820.
22. Hake RR. Interactive-engagement vs. traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am J Phys* 66: 64–74, 1998. doi:10.1119/1.18809.
23. Handelsman J, Miller S, Pfund C. *Scientific Teaching*. New York: Freeman, 2007.
24. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. *Commun Methods Meas* 1: 77–89, 2007. doi:10.1080/19312450709336664.
25. Henderson C, Dancy M. Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Phys Rev Spec Top Phys Educ Res* 3: 020102, 2007. doi:10.1103/PhysRevSTPER.3.020102.
26. Henderson C, Dancy M, Niewiadomska-Bugaj M. The use of research-based instructional strategies in introductory physics: Where do faculty leave the innovation-decision process? *Phys Rev Spec Top Phys Educ Res* 8: 020104, 2012. doi:10.1103/PhysRevSTPER.8.020104.
27. Hestenes D, Wells M, Swackhamer G. Force concept inventory. *Phys Teach* 30: 141–158, 1992. doi:10.1119/1.2343497.
28. Karpicke JD, Roediger HL III. The critical importance of retrieval for learning. *Science* 319: 966–968, 2008. doi:10.1126/science.1152408.
29. Knight JK, Wood WB. Teaching more by lecturing less. *Cell Biol Educ* 4: 298–310, 2005. doi:10.1187/05-06-0082.
30. Kornell N, Bjork RA. The promise and perils of self-regulated study. *Psychon Bull Rev* 14: 219–224, 2007. doi:10.3758/BF03194055.
31. Kruger J, Dunning D. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Psychology (Irvine)* 1: 30–46, 2009.
32. Lasry N, Mazur E, Watkins J. Peer instruction: From Harvard to the two-year college. *Am J Phys* 76: 1066–1069, 2008. doi:10.1119/1.2978182.
33. Mazur E. *Peer Instruction: A User's Manual*. Upper Saddle River, NJ: Prentice Hall, 1997.
34. Mervis J. Grand challenge: undergraduate teaching. Transformation is possible if a university really cares. *Science* 340: 292–296, 2013. doi:10.1126/science.340.6130.292.
35. National Research Council. *How People Learn: Brain, Mind, Experience and School*. Washington, DC: National Academies, 2000.
36. National Research Council. *Committee on the Status, Contributions, and Future Directions of Discipline-Based Education Research. Board on Science Education, Division of Behavioral and Social Sciences and Education. Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*, edited by Singer SR, Nielsen NR, Schweingruber HA. Washington, DC: National Academies, 2012.
37. Norman GR, Schmidt HG. The psychological basis of problem-based learning: a review of the evidence. *Acad Med* 67: 557–565, 1992. doi:10.1097/00001888-199209000-00002.
38. Perez KE, Hiatt A, Davis GK, Trujillo C, French DP, Terry M, Price RM. The EvoDevoCI: a concept inventory for gauging students' understanding of evolutionary developmental biology. *CBE Life Sci Educ* 12: 665–675, 2013. doi:10.1187/cbe.13-04-0079.
39. Prince M. Does active learning work? A review of the research. *J Eng Educ* 93: 223–231, 2004. doi:10.1002/j.2168-9830.2004.tb00809.x.
40. Reeves TC. How do you know they are learning?: the importance of alignment in higher education. *Int J Learn Technol* 2: 294–309, 2006. doi:10.1504/IJLT.2006.011336.
41. Seymour E. Tracking the process of change in us undergraduate education in science, mathematics, engineering, and technology. *Sci Educ* 86: 79–105, 2002. doi:10.1002/sci.1044.
- 41a. Semsar K, Casagrand J. Bloom's dichotomous key: a new tool for evaluating the cognitive difficulty of assessments. *Adv Physiol Educ* 41: 170–177, 2017.
42. Silverthorn DU. Teaching and learning in the interactive classroom. *Adv Physiol Educ* 30: 135–140, 2006. doi:10.1152/advan.00087.2006.
43. Silverthorn DU, Thorn PM, Svinicki MD. It's difficult to change the way we teach: lessons from the Integrative Themes in Physiology curriculum module project. *Adv Physiol Educ* 30: 204–214, 2006. doi:10.1152/advan.00064.2006.
44. Simon B, Taylor J. What is the value of course-specific learning goals? *J Coll Sci Teach* (November/December): 52–57, 2009.
45. Smith MK, Perkins KK. "At the end of my course, students should be able to . . .": The benefits of creating and using effective learning goals. *Microbiol Aust* (March): 35–37, 2010.
46. Smith MK, Wood WB, Knight JK. The Genetics Concept Assessment: a new concept inventory for gauging student understanding of genetics. *CBE Life Sci Educ* 7: 422–430, 2008. doi:10.1187/cbe.08-08-0045.
47. Smith MK, Wood WB, Krauter K, Knight JK. Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE Life Sci Educ* 10: 55–63, 2011. doi:10.1187/cbe.10-08-0101.
48. Sundberg MD. Assessing student learning. *Cell Biol Educ* 1: 11–15, 2002. doi:10.1187/cbe.02-03-0007.
49. Tagg J. Why does the faculty resist change? *Change* 44: 6–15, 2012. doi:10.1080/00091383.2012.635987.
50. Vicens Q, Caspersen ME. Getting more scientists to revamp teaching. *J Coll Sci Teach* 43: 22–27, 2014. doi:10.2505/4/jcst14_043_05_22.
51. Wieman C, Deslauriers L, Gilley B. Use of research-based instructional strategies: How to avoid faculty quitting. *Phys Rev Spec Top Phys Educ Res* 9: 023102, 2013. doi:10.1103/PhysRevSTPER.9.023102.
52. Wieman C, Perkins K, Gilbert S. Transforming Science Education at Large Research Universities: A Case Study in Progress. *Change* (March/April): 7–14, 2010.
53. Wood WB. Innovations in teaching undergraduate biology and why we need them. *Annu Rev Cell Dev Biol* 25: 93–112, 2009. doi:10.1146/annurev.cellbio.24.110707.175306.
54. Zheng AY, Lawhorn JK, Lumley T, Freeman S. Assessment. Application of Bloom's taxonomy debunks the "MCAT myth". *Science* 319: 414–415, 2008. doi:10.1126/science.1147852.