JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

Alcohol Pharmacology Education Partnership: Using Chemistry and Biology Concepts To Educate High School Students about Alcohol

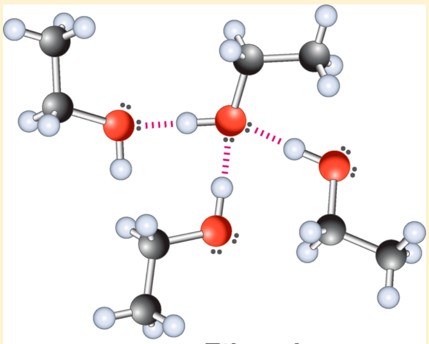
Elizabeth A. Godin,*[,†] Nicole Kwiek,[†,⊥] Suzanne S. Sikes,[†,¶] Myra J. Halpin,[‡] Carolyn A. Weinbaum,[†] Lane F. Burgette,[§,∥] Jerome P. Reiter,[§] and Rochelle D. Schwartz-Bloom[†]

[†]Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, North Carolina 27708 United States

[‡]Department of Chemistry, North Carolina School of Science and Mathematics, Durham, North Carolina 27705 United States

[§]Department of Statistical Science, Duke University, Durham, North Carolina 27708 United States

Supporting Information



ABSTRACT: We developed the Alcohol Pharmacology Education Partnership (APEP), a set of modules designed to integrate a topic of interest (alcohol) with concepts in chemistry and biology for high school students. Chemistry and biology teachers (n = 156) were recruited nationally to field-test APEP in a controlled study. Teachers obtained professional development either at a conference-based workshop (NSTA or NCSTA) or via distance learning to learn how to incorporate the APEP modules into their teaching. They field-tested the modules in their classes during the following year. Teacher knowledge of chemistry and biology concepts increased significantly following professional development, and was maintained for at least a year. Their students (n = 14 014) demonstrated significantly higher scores when assessed for knowledge of both basic and advanced chemistry and biology concepts compared to students not using APEP modules in their classes the previous year. Higher scores were achieved as the number of modules used increased. These findings are consistent with our previous studies, demonstrating higher scores in chemistry and biology after students use modules that integrate topics interesting to them, such as drugs (the Pharmacology Education Partnership).

KEYWORDS: *High School/Introductory Chemistry, Chemical Education Research, Interdisciplinary/Multidisciplinary, Distance Learning/Self Instruction, Inquiry-Based/Discovery Learning, Testing/Assessment, Alcohols, Drugs/Pharmaceuticals, Enzymes, Oxidation/Reduction*

FEATURE: Chemical Education Research

■ INTRODUCTION

It is no surprise that alcohol, as a topic, is inherently interesting to high school students; it is one of the most commonly used drugs by this population. According to the Center for Disease Control's Youth Risk Behavior Surveillance System, 71% of high school students have consumed alcohol on at least one day in their life.[1] It is even more alarming that 57% of 12th graders have been drunk at least once. Clearly, there is a need to educate students on the dangers of drinking alcohol, especially at a young age. Typically, substance abuse topics are covered in health education courses—not courses in chemistry or biology. Yet, basic principles of chemistry and biology are crucial in understanding our bodies, health, and disease.

Educators can take advantage of the fact that the topic of alcohol interests high school students. Promoting interest has been shown to be an important factor in helping motivate students to learn and increase achievement.[2] Moreover, Linnenbrink-Garcia, et al.[3] have shown that using topics that connect to real life in the classroom is one way to increase interest and motivation in science learning. In fact, high school students indicate that the topics of most interest to them in their science classes include drugs, disease, and the environment.[4] Moreover, students with relatively low expectations for success in science display more interest and perform better in science when they are asked to connect the relevance of their science topics in class to their lives.[5] There have been relatively few studies that assess how specific topics affect student achievement in chemistry and biology. One study that used the *Chemistry in the Community* curriculum (ChemComm),[6] which focuses on areas such as the environment, industry, food, and health, demonstrated increased achievement in a small group of high school chemistry students.[7] In several larger studies, we have shown that curricula focused on pharmacology topics such as drugs of abuse can markedly improve high school student achievement in chemistry and biology.[8−10] The very nature of pharmacology, which integrates basic principles of chemistry and biology to uncover the mechanisms by which drugs and chemicals affect organisms, lends itself to a useful approach for providing relevance and context—especially for teenagers. In our previous studies, teachers used a series of modules in their classrooms as part of our program called the Pharmacology Education Partnership (PEP).[8−10] The PEP modules covered various drug-related topics with catchy titles such as "Acids, Bases, and Cocaine Addicts" and "Steroids and Athletes—Genes Work Overtime". These modules are available free to the public and are found online.[11]

The PEP studies provided the basis for the current study, which brings the subject of alcohol pharmacology to high school teaching. The teachers who participated in PEP suggested that the subject of alcohol was of great interest to their students and thought that a similar program devoted to alcohol as a context for learning chemistry and biology would be well received. Moreover, teachers mentioned that there is considerable misconception among their students about the chemistry and biology of alcohol. For example, many adolescents believe that drinking alcohol once an hour can avoid legal intoxication, yet a basic understanding of chemistry and biology can help to correct this misconception. In addition, teachers reported that students have misconceptions of basic chemical and biological principles as well, such as the meaning of equilibrium and the relationship between genes and proteins. Research in areas of cognitive psychology reveals that students' prior knowledge (including misconceptions) and their prior experiences are crucial elements in formulating their own understanding of phenomena (constructivism theory).[12] Thus, a program that can challenge prior knowledge and correct misconceptions about chemistry and biology that relate to a topic such as alcohol may help students construct meaning from their everyday life and better their learning in these sciences.

To implement such a program, professional development at the teacher level is just as important. Effective professional development aims to improve teacher knowledge, teaching practices, and ultimately student achievement. As compiled by Supovitz and Turner,[13] high-quality professional development should include at least the following:

- Inquiry-based approaches and active learning.
- An intensive and sustained program.
- A mechanism to engage teachers in concrete teaching tasks that they can use in their own classrooms.
- A focus on subject-matter knowledge and content skills.

We addressed each of these features in our three previous PEP studies,[8−10] which delivered professional development at (i) a five-day residential workshop, (ii) a one-day workshop at a national meeting or (iii) a six-hour workshop via distance learning (two hours per week). Regardless of the workshop format, teacher content knowledge in chemistry and biology increased and was maintained for at least a year. Additionally, in each case, there was greater student achievement compared to students who did not have the PEP modules in their classes. Thus, our findings support research indicating that rich content, rather than the duration of professional development, is associated with student learning.[14]

For this study, we recruited a new set of teachers from the United States to participate in the Alcohol Pharmacology Education Partnership (APEP). We used the more cost-effective forms of professional development, a full-day workshop onsite at an annual conference and an interactive three-session workshop delivered by distance learning technology. Each of the APEP modules that we developed addresses several basic principles of chemistry and biology. Additionally, the modules capitalize on the natural curiosity that teens typically have concerning the subject of alcohol. As we demonstrated previously using PEP,[8−10] students scored higher when tested for knowledge of chemistry and biology concepts after using the APEP modules than students who did not use the modules.

■ INTERVENTION

Participants

Participating Teachers. High school chemistry and biology teachers were recruited from across the United States by placing announcements in the National Science Teachers Association (NSTA) newspaper, *NSTA Reports*. Ultimately, 156 teachers participated in the study. Details of the teacher demographics are provided in the Supporting Information, Table S1. In brief, 50% of teachers were of biology or chemistry, 88% were from public schools, 30% were teaching in schools with at least a 40% population of minorities, and 67% were from urban or suburban schools.

Teachers selected to attend six hours of APEP professional development at either a full-day conference-based workshop conducted concurrently with the NSTA or North Carolina Science Teachers Association (NCSTA) annual meetings or a series of distance learning (DL) workshops equal in length to the residential sessions. Of the 156 participating teachers, 100 attended the conference-based workshops and 56 attended the DL workshops.

Participating Students. Teachers' students (n = 14 014) who participated in the study were from chemistry and biology classes in grades 9−12 representing schools from urban, suburban, and rural districts. The demographics of students participating in the study are shown in Table 1. During the first year of the study, all students served as the control group (before teachers attended any APEP professional development); the following year, a second set of students served as the experimental group. At the start of the second year, teachers attended the professional development workshop and then field-tested the APEP curriculum in their classrooms over the ensuing school year. There was no indication that demographics of the school student populations differed systemati-

Table 1. Demographics of Students in Classes of Participating Teachers

demographic variable	students, % (n = 14 014)
gender	
male	56.1
female	43.9
class year	
freshman	15.9
sophomore	37.0
junior	32.4
senior	14.7
race and ethnicity	
Caucasian	68.0
Asian	7.2
Black	14.4
Native American	2.0
Hispanic	8.3
course	
Chemistry 1	45.2
Biology 1	35.1
Chemistry 2 or AP[a]	4.6
Biology 2 or AP[a]	15.1

[a]AP designates an advanced placement course.

Table 2. APEP Module Chemistry and Biology Content

module title	chemistry content	biology content	other content	NGSS disciplinary core ideas
gender matters	oxidation−reduction; enzymes (catalysts); solubility; molecular structure; polarity	cell structure; cell types; membrane transport; anatomy; circulatory system	addiction biology; gender issues; algebra	PS1A, PS1B; LS1A, LS1B, LS3A, LS3B
ABCs of intoxication	chemical structure and bonds; enzymes	membrane transport; passive diffusion; cell structure; brain anatomy; DNA	algebra; genetics	PS1A, PS1B; LS1A, LS1B, LS3A, LS3B, LS4B
alcohol, cell suicide, and the adolescent brain	atomic structure; enzymes (catalysts)	cell cycle; protein synthesis; cell death; brain anatomy; cell structure; neurons	apoptosis	PS1A, PS1B, PS2B, PS4B; LS1A, LS1B, LS3A, LS3B
alcohol and the breathalyzer test	redox reactions; solubility; intermolecular forces; equilibrium; chemical equations	anatomy; circulatory system; diffusion	physiology	PS1A, PS1B; LS1A

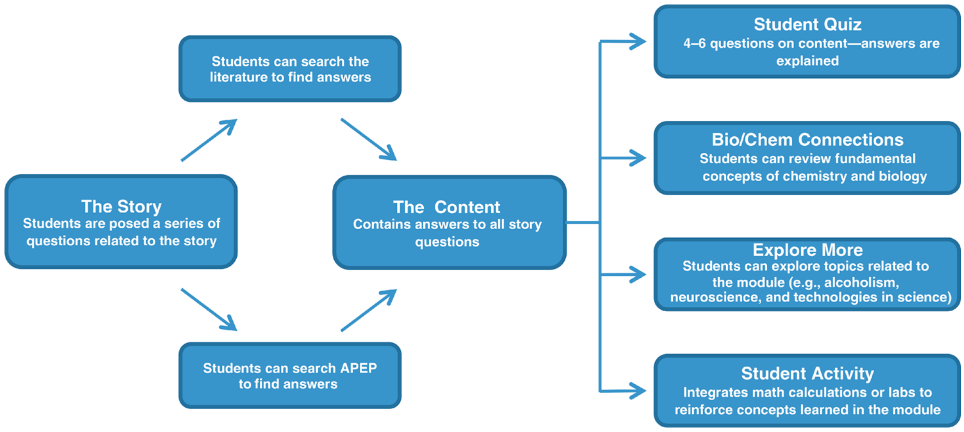


Figure 1. General structure of the APEP modules.

cally from year to year. Thus, each teacher served as his or her own control.

Professional Development for APEP Teachers

All teachers enrolled in the study received professional development provided by three of us (R.D.S.-B., S.S.S., and M.J.H.) either at the annual NSTA or NCSTA meeting or via DL. See Supporting Information Table S2 for the regional demographics of broadcast sites for the DL workshops. The DL workshops were conducted using two-way live audio and video plus high-speed data to allow interaction between instructors and teachers as we reported previously.[10] This two-way system allowed connections to four sites simultaneously and was broadcast from the North Carolina School for Science and Mathematics in Durham, NC. (An example of the DL broadcast is available.[15])

The formats for both the conference-based and DL workshops were based on our previous PEP studies.[9,10] Central chemistry concepts (e.g., dynamic equilibrium, solubility, molecule polarity, enzymes, catalysts) and biology concepts (e.g., cell structure and function, transport of molecules across membranes, protein synthesis) were presented within the context of four APEP modules developed by the authors (see below) to be used in field-testing in the teachers' classes the following year. Thus, during the workshops, teachers not only were able to review each of these concepts and learn how to incorporate them within the APEP modules but also tackled one of the APEP modules themselves, working in groups, just as their students would. They read the opening story, researched answers to the inquiries built into the module, discussed their findings with their partners, and then presented their answers to all of the teachers.

The teachers also participated in one of the hands-on laboratory exercises contained in the APEP curriculum. Teachers prepared a chemical Breathalyzer test that can be adapted for standard or advanced placement (AP) courses. Teachers in the DL course also participated in the laboratory exercise during the live broadcast with materials provided to them ahead of time.

APEP Modules

The team of authors created four APEP modules to be field-tested in classrooms during the study. The modules were supplied in print, DVD, and online versions. The four modules integrate chemistry and biology concepts within the context of alcohol-related pharmacology topics that have relevance to high school students. The modules are listed in Table 2. Each module contains the following elements:

- An overview containing a description of the module with the associated education standards.
- A set of learning objectives.
- A student handout that contains the story followed by a set of questions that students must answer (similar to a problem in problem-based learning).
- The content—the bulk of the module that contains all the answers to the student questions, including animations.
- A student quiz with 4−6 questions about the chemistry and biology concepts relevant to the module (and answers provided).
- Activities (integrating math calculations or lab activities with the module content).
- Resources (a list of references and URL links).

In addition to these module-specific items, there are several other resources that address all the modules. These include:

- Biology and Chemistry Connections—a table showing the specific chemistry and biology concepts that are covered in each module and links to a review of specific chemistry and biology concepts normally covered in high school courses.
- Explore More, a set of pages (also linked within the modules) that cover topics related to the modules such as alcoholism and neuroscience, and technologies such as magnetic resonance imaging..
- A glossary.
- List of the Next Generation Science Standards and former National Science Education Standards.

The module content is aligned with the former National Science Education Standards[16] and with the new Next Generation Science Standards (NGSS)[17] (see Table 2). In addition, the APEP module content addresses typical teen misconceptions and myths related to alcohol consumption. These include misconceptions such as (i) drinking one drink an hour will not get you drunk, (ii) you can get rid of alcohol by urinating it out, and (iii) beer and wine are less intoxicating than liquor. The modules were developed to foster student engagement in the form of a story, similar to a problem-based or case-based learning approach. Working in groups, students read the story contained in the student handout, which is followed by a series of questions relating to the story. They must do some research to answer the questions. It is up to the teacher as to whether the students can use Web-based literature to find the answers, or use the APEP module content pages to find the answers. A student self-assessment quiz is included within each module with an explanation of the correct or incorrect answers. In addition, the modules contain activities, many of which include mathematical calculations. The general structure of the APEP modules is shown in Figure 1 and the story for Module 4 is provided in the Supporting Information. The APEP modules may be accessed online[15] (both a teacher and student version).

APEP Module Field Test

Following the training, teachers were instructed to field test the four APEP modules in their classrooms over the year. We did not prescribe a specific approach for teachers to use the APEP modules; instead, we asked teachers to report how they incorporated the content into their courses. We encouraged them to use as many modules as possible. Previously, we showed that the more PEP modules used, the greater the beneficial effect on student scores in both basic and advanced science knowledge.[8−10]

Assessments for Teachers

First, we conducted a summative evaluation to determine the attitudes of teachers about the quality of the workshop. There were three strands within the survey: content, teaching approaches, and format of the workshop. Items were assessed using a five-point Likert-type scale, followed by several open-ended questions pertaining to what teachers liked most and least about the workshop. The evaluation items can be found in the Supporting Information, Table S3.

To determine the effectiveness of the workshop on teachers' knowledge gain and retention after one year, we administered a short test consisting of 20 multiple-choice questions that addressed the chemistry and biology content listed in Table 2. The test was administered to the teachers at the beginning of the workshop (pretest), at the end of the workshop (post-test), and again at the end of the year (one-year follow-up) without prior notification. Sample items are included in the Supporting Information.

Assessments for Students

At the end of the school year after field-testing the APEP modules, we sent the teachers a multiple-choice content test similar to that used in our previous PEP studies[8−10] to give to their students; the tests were unannounced. The tests were constructed by the authors, with input from high school chemistry and biology teachers at the North Carolina School of Science and Math, where one of us (M.J.H.) teaches. The test comprised two parts, a basic knowledge and an advanced knowledge section. The basic test consisted of 20 questions (11 chemistry and 9 biology) similar to those found in first-year chemistry and biology textbooks (see the supporting online material for examples). The multiple choice questions assessed student knowledge of concepts in chemistry and biology according to the framework provided by the 1996 National Assessment of Educational Progress science test.[18] Following the 20 basic knowledge questions, there were 10 questions that were specific for the new knowledge about alcohol in the context of chemistry and biology (advanced knowledge). These questions assessed concepts not normally taught in the standard curriculum (see the Supporting Information for examples).

To establish reliability of the assessment instrument, we conducted a test−retest with 58 high school students in chemistry and biology classes (unrelated to the study). The tests were administered 10 days apart. An intraclass correlation of scores between the two tests generated a reliability coefficient of 0.85 (95% confidence interval 0.73−0.92). Reliability coefficients 0.81−1.0 are considered substantial.[19]

To establish validity of the assessment instrument, we asked a team of nine high school chemistry and biology teachers (not related to any of the authors) to rate each of the questions as to whether it was relevant and appropriate to their courses. For the biology and chemistry basic questions, eight of nine teachers rated at least 80% and 90%, respectively, of the questions as relevant and appropriate. For the advanced questions, seven of nine teachers rated at least 80% of the biology questions as relevant and appropriate, and all of the teachers rated at least 80% of the chemistry questions as relevant and appropriate.

We posited that students with different backgrounds could score differently on average on the APEP tests. Therefore, we obtained demographic information from the students regarding these demographic parameters: gender, race and ethnicity, year in high school (i.e., 9th−12th grade), course type (i.e., chemistry or biology), and course level (i.e., first-year, second-year, or AP). The demographic representation of students within classes of teachers who administered the APEP tests is presented in Table 1.

Data Analysis: Statistical Model

We compiled the percent correct scores of the 14 014 students on both the basic and advanced tests. To estimate the effects of the modules, we used logistic regression models with random effects for the teachers (which is a type of multilevel model[20,21]), adjusting for demographic characteristics (see Table 1), and number of modules as a series of indicator variables. The outcome variables are the number of correct answers out of 20 questions on the basic test and the number of correct answers out of 10 questions on the advanced test. For simplicity, we analyzed each outcome independently. In the

Table 3. Effect of Professional Development on Teacher Knowledge

teacher assessment[a]	DL workshop (n = 29) mean (SD) for percentage of correct answers	DL workshop effect size, Cohen's d values	DL workshop p values[b]	conference workshop (n = 82) mean (SD) for percentage of correct answers	conference workshop effect size, Cohen's d values	conference workshop p values[b]
pretest	59 (16.0)			61 (18.2)		
post-test	70 (17.8)	0.84	<0.01	74 (16.4)	0.79	<0.01
one-year follow-up	70 (21.7)	0.72	<0.01	78 (16.4)	1.04	<0.01

[a]Data are from teachers who participated in all three testing modalities: pretest, post-test, and one-year follow up. [b]The p values are for comparisons to the pretest. There was no significant difference between the post-test and one-year follow-up test for either workshop.

binomial regression models, we used both student-level and teacher-level random effects. Student-level random effects for students taught by the same teacher are centered around their teacher-specific mean, which enables us to account for nesting of students within teachers. Thus, the random effects account for the correlations among outcomes of students taught by the same teacher. Additional details about the models can be found in the Supporting Information.

Exemption from Human Subjects Review

This research received an IRB approval of exemption from human research subjects according to federal rules (45 CFR 46.101(b)(1)) and is not subject to the Privacy Rule (HIPPA) (45 CFR 164.500(a)).

■ RESULTS

Teacher Content Knowledge

To assess the effect of providing the pharmacology-based professional development workshop on teacher knowledge of basic chemistry and biology concepts, teachers were assessed for content knowledge at the beginning of the workshop (pretest), at the end of the workshop (post-test), and one year following the workshop (one-year follow-up). Teacher scores are summarized in Table 3. Based on separate repeated measures ANOVAs for each workshop, teacher scores differed significantly among the three tests for the DL workshop, F (1.852, 53.70) = 11.79 and $p < 0.0001$, and for the conference workshop, F (1.78, 144.3) = 42.33 and $p < 0.0001$. All score distributions were approximately distributed as Gaussian, with a Geisser−Greenhouse epsilon of 0.93 and 0.89, respectively. The ANOVA effect sizes (η^2) were 0.05 and 0.16, respectively. These effect sizes range from medium to large according to Cohen.[22] Comparisons between the pre- and post-tests revealed that average scores on both the post-test and the one-year follow-up test were significantly higher than the pretest averages (Tukey's multiple comparison test). There was no difference between the scores in the post-test and the one-year follow-up test for both workshops, indicating persistence of the knowledge gain. The effect sizes (Cohen's d) for knowledge gains were in the moderate to large range.

We also assessed whether teachers' knowledge gain was similar on the biology and chemistry questions depending on whether the teacher taught biology or chemistry. There were no statistically significant differences in knowledge gains over the year for any comparison with the exception of chemistry teachers who attended the conference workshop. Their average gain in scores on the biology questions over the year was considerably larger than their gain in scores on the chemistry questions (mean gain in scores ± SD was 31 ± 32 versus 11 ± 26 percentage points, respectively (n = 31); $p < 0.05$, paired Student's t test).

Teacher Evaluation of Conference-Based and Distance Learning Workshops

At the conclusion of the workshop, teachers provided an evaluation of the workshop approach and content delivery. Overall, the teachers provided a quite positive review of the workshops. A summary of their evaluation scores is provided in the Supporting Information, Table S3.

Student Achievement in Chemistry and Biology

Teachers participating in the workshop were instructed to use as many APEP modules as possible (i.e., four) in their classes throughout the year. Although there was no set prescription for implementing the modules, we did collect information on how the teachers used the modules in their classes (Table 4). Of the 426 classes in which modules were used, the most common format was to cover the module content over several class periods (49% of classes).

Table 4. Implementation of the APEP Modules

module implementation method	classes, %
incorporated parts of module (subjects) into teaching throughout the course	11
gave lectures on module material during a single class session	22
gave lectures on module material over several class sessions	49
assigned modules as homework or group projects	13
other (e.g., students worked online)	6

Teachers administered the APEP tests, unannounced, at the end of their courses. The results revealed that the use of the APEP modules was a significant predictor of greater student achievement on both the basic and advanced knowledge tests, as shown in Figure 2 and in the Supporting Information. Using more modules resulted in progressively higher scores on both tests. The average scores in the control teachers' classrooms using zero modules did not differ practically or significantly from the average scores in classrooms of teachers in the experimental group who did not use any APEP modules; hence, we used a common indicator variable for zero APEP modules for these two sets of students. The regression results suggested that usage of modules is associated with higher test scores. For example, a typical student in a class that used all four modules had 1.21 times the odds of answering basic knowledge questions correctly compared to a typical student in a class that used no modules. Not surprisingly, the strongest predictor of higher basic and advanced scores was being in an advanced biology or chemistry course. A typical Biology 2 student had 1.58 times the odds of answering a basic knowledge question correctly compared to a typical Biology 1 student, and a typical Chemistry 2 student had 1.84 times the odds of answering correctly compared to a typical Biology 1 student. For the basic knowledge questions, students in Chemistry 1 had a small

dx.doi.org/10.1021/ed4000958 | J. Chem. Educ. XXXX, XXX, XXX−XXX

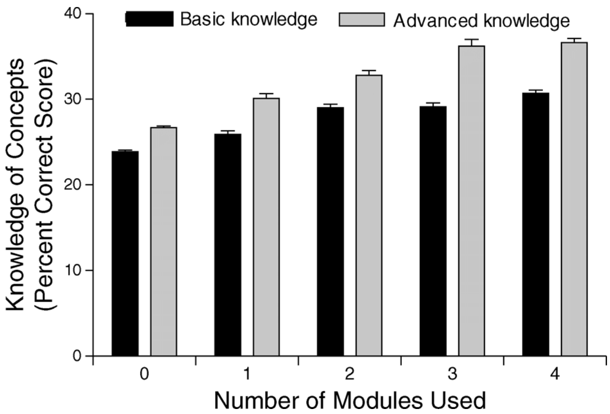


Figure 2. Student knowledge assessed at the end of the course (chemistry or biology) in which APEP was administered. Student basic and advanced knowledge scores were higher as more modules were used in their classroom. Data represent the mean ± SEM scores (unadjusted) from students in classes of teachers enrolled in the APEP study. The zero module data includes students from the control year and teachers who taught zero modules in the experimental year. Use of modules was a significant predictor of higher scores (logistic regression; n = 14 014 students; see the Supporting Information, Tables S4 and S5).

increase in odds of answering correctly compared to Biology 1 students. Because most chemistry students had already had biology, this might be expected. Several student demographic characteristics were also statistically significant predictors in both models, though with smaller coefficients than the class type. All results are in the regression tables in the Supporting Information, Tables S4 and S5.

Students in the APEP study were enrolled in either chemistry or biology classes. We also analyzed subsets of questions and students separately (e.g., the performance of biology students on chemistry questions, the performance of chemistry students on chemistry questions, etc.). In general, using more modules was associated with higher test scores, preserving the dose-response relationship we see in the overall analysis (Table 5). However, the higher chemistry student scores on biology and chemistry questions remained relatively flat when one, two, and three modules were used compared to no modules. When four

Table 5. Chemistry and Biology Knowledge Increases for Both Chemistry and Biology Students

		correct answers, % ± SE[a]	
students	modules used	chemistry	biology
chemistry	0	25.5 ± 0.2	27.0 ± 0.2
	1	28.8 ± 0.6	30.3 ± 0.7
	2	29.0 ± 0.7	30.4 ± 0.6
	3	28.0 ± 0.8	30.7 ± 0.8
	4	31.0 ± 0.5	32.7 ± 0.5
biology	0	19.2 ± 0.2	27.4 ± 0.2
	1	19.5 ± 0.5	30.0 ± 0.7
	2	24.7 ± 0.6	35.9 ± 0.7
	3	27.3 ± 0.7[b]	37.6 ± 0.9
	4	28.7 ± 0.6[b]	37.6 ± 0.7

[a]Data represent the mean percentage correct scores ± SE (unadjusted). Use of modules was a significant predictor of higher scores (logistic regression; n = 14 014 students). [b]Students in biology class whose teachers used three or four APEP modules achieved chemistry scores similar to the baseline scores of students in chemistry.

modules were used, there was an additional increase in scores. Thus, the steady increase in scores from biology students when using one, two, and three modules may have accounted for most of the higher scores as reported in the overall analysis (Figure 2). Interestingly, students in biology class whose teachers used three or four APEP modules achieved chemistry scores similar to the baseline scores of students in chemistry (see Table 5).

DISCUSSION

In this study, we demonstrate that implementing pharmacology-based science education (using alcohol as the context) in high school classrooms increases student and teacher achievement in chemistry and biology. These findings reproduce the findings obtained in our three previous PEP studies that focused on drugs.[8−10] In the present study, we found up to 10 percentage-point increases in achievement (compared to no module use) when students were administered unannounced tests. Even greater increases can be achieved when using up to six modules, as we showed in a previous PEP study that focused on drugs.[10] In the APEP study, all of the modules focused on one drug (alcohol), so it is possible that limiting the subject matter may have contributed somewhat to a smaller effect compared to our previous studies that incorporated a variety of drug topics.[8−10] The format of the APEP curriculum is similar to the PEP curriculum, so it would be easy to combine the two, thereby increasing the number of modules (to 10) that could be used to teach chemistry and biology concepts.

The content within APEP was designed to address the 1996 National Science Education Standards.[16] Since then, the Standards have been revised, recommending an integrated, research-based curriculum.[23] The integration of chemistry, biology, and math concepts in APEP addresses several components of the new NGSS,[17] including scientific practices and the integration of disciplinary core ideas. We surmise that integration of chemistry and biology may have been successful in our study because students in biology, who typically take chemistry a year after biology, were able to increase their chemistry scores to the same level as that achieved by control chemistry students just by using at least three APEP modules.

The APEP curriculum has several features that may have contributed to the positive outcome, including the following:

- The relevance of the various alcohol-focused stories (and the content) to the high school population.
- An inquiry-based approach to learn the content.
- Content that addresses myths and misconceptions about alcohol.
- Repetition of core concepts among the modules.

Each of these features are important elements in teaching strategies for effective learning.[24−26] Additional studies would be required to isolate these variables to determine which is the greatest predictor of higher scores. As mentioned above, we have shown in a previous PEP study[10] that use of six modules predicts even higher scores in chemistry and biology than reported here, supporting the importance of dose. However, we suggest that it is the combination of these features that makes the APEP curriculum complete, as it includes elements of best practices in curriculum design and teaching.[24−26]

In terms of professional development, both on-site conference-based workshops and the distance learning workshops helped teachers improve their knowledge in chemistry and biology and maintain that improvement for at least a year.

However, merely attending the professional development workshop was not sufficient for increased student achievement. Student scores of the control teachers did not differ from the experimental teachers who attended the workshop but did not use the APEP modules. This is important because there is a plethora of professional development activities for teachers—many providing continuing education credit—that assume the experience will improve teaching. Our findings suggest that teachers must use what they have learned in the workshop along with the curriculum to help students learn science better.

LIMITATIONS

Although the results presented here replicate what we've found in our previous studies,[8−10] there are several limitations. First, although teachers used the modules in a variety of ways, we do not know which of the implementation methods was a better predictor for the stepwise increase in scores. Because several teachers used multiple implementation methods in their classes at different times, it was not realistic to determine whether any method was more effective than another. However, presenting the module material over several classes was the most prevalent implementation format (~50% of classes used this format), so it is reasonable to conclude that at least this particular method may have contributed significantly to the increasing scores. In contrast, only 6% of classes worked on the APEP modules online, so it is unlikely that this method alone contributed significantly to the positive findings. In addition, we do not know whether any specific module was important in explaining the increasing scores. As there was considerable overlap of concepts from module to module, it is possible that repetition of concepts rather than any single module may have contributed to the increasing scores with increased module use.

Another limitation of the study is the possibility that teachers using the modules were "teaching to the test". We believe that this is unlikely because the questions for the basic test were not directed specifically to the module content but rather to the principles associated with the module content. Although the advanced test questions did address specific module content, many required the students to apply what they learned from the modules to a new situation. Nevertheless, this limitation is still possible for the advanced questions.

Third, the improvements in the scores associated with module usage may not be caused by the modules themselves; other variables that we did not control could explain (at least in part) the higher scores. For example, suppose that a large majority of the students who used the APEP modules have higher academic motivation and abilities than those who did not use the modules, and this difference, rather than beneficial effects of learning with the modules, accounts for their higher average test scores. However, it is unlikely that a systematic disparity in student abilities would be present because the disparities would also need to be present in a dose-response manner (as were the higher scores), and classes of the same type taught by the same teacher were likely to be similar in terms of prematriculation abilities. Nevertheless, such confounders are always possible given that we did not control for student ability.

CONCLUSIONS

This is the fourth study in which we have demonstrated that students who use problem-based units addressing relevant topics to their lives (e.g., alcohol and drugs) score better in chemistry and biology compared to students using the standard curriculum. In total, we have now tested 27 841 students using both randomized[8] and nonrandomized[9,10] controlled designs that generate the same results. The implementation of controlled research designs can provide teachers with meaningful information about curricular materials and teaching strategies most likely to help their students learn chemistry and biology better. Additionally, a major goal in science education is to help students become critical thinkers and make good decisions about their daily lives. It remains to be determined whether learning the chemical and biological principles underlying alcohol's actions will impact students' decisions about alcohol use.

The APEP modules can be accessed online free of charge for interactive use or downloaded as PDFs directly from our Web site.[15] A fifth module addressing fetal alcohol spectrum disorders has been added since the APEP study was completed; however, it was not used for any data collection.

ASSOCIATED CONTENT

Supporting Information

Tables; sample teacher and student assessment items; statistical methods. This material is available via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E. A. Godin. E-mail: elizabeth.godin@duke.edu.

Present Addresses

[⊥]Department of Pharmacology, Ohio State University, Columbus, Ohio 43210, United States.
[¶]Department of Biology, Woods Charter School, Chapel Hill, North Carolina 27516, United States.
[‖]RAND Corporation, Arlington, Virginia 22202, United States.

Notes

Disclaimer

Research reported in this publication was supported by the National Institute of Alcohol Abuse and Alcoholism of the National Institutes of Health under Award Number AA15429. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by NIH/NIAAA grant No. AA15429.

REFERENCES

(1) Centers for Disease Control and Prevention. Youth Risk Behavior Surveillance—United States, 2011. *Morbidity and Mortality Weekly Report* **2012**, *61* (4), 1−162.

(2) Sandoval, J. Teaching in Subject Matter Areas: Science. *Annu. Rev. Psychol.* **1995**, *46*, 355−374.

(3) Linnenbrink-Garcia, L.; Patall, E. A.; Messersmith, E. E. Antecedents and Consequences of Situational Interest. *Br. J. Educ. Psychol.* **2013**, *83* (4), 591−614.

(4) Jenkins, E. W.; Nelson, N. W. Important But Not for Me: Students' Attitudes towards Secondary School Science in England. *Res. Sci. Technol. Educ.* **2005**, *23*, 41−57.

(5) Hulleman, C. S.; Harackiewicz, J. M. Promoting Interest and Performance in High School Science Class. *Science.* **2009**, *326*, 1410−1412.

(6) American Chemical Society. *Chemistry in the Community*, 5th ed.; W. H. Freeman: New York, 2006.
(7) Winther, A. A.; Volk, T. L. Comparing Achievement of Inner-City High School Students in Traditional versus STS-Based Chemistry Concepts. *J. Chem. Educ.* **1994**, *71* (6), 501−505.
(8) Schwartz-Bloom, R. D.; Halpin, M. J. Integrating Pharmacology Topics in High School Biology and Chemistry Classes Improves Performance. *J. Res. Sci. Teach.* **2003**, *40* (9), 922−993.
(9) Kwiek, N. C.; Halpin, M. J.; Reiter, J. P.; Hoeffler, L. A.; Schwartz-Bloom, R. D. Pharmacology in the High School Classroom. *Science* **2007**, *317*, 1871−1872.
(10) Schwartz-Bloom, R. D.; Halpin, M. J.; Reiter, J. P. Teaching High School Chemistry in the Context of Pharmacology Helps Both Teachers and Students Learn. *J. Chem. Educ.* **2011**, *88*, 744−750.
(11) Pharmacology Education Partnership start page. http://www.thepepproject.net (accessed Jan 2014).
(12) Mintzes, J. J.; Wandersee, J. H.; Novak, J. D. Teaching Science for Understanding: A Human Constructivist View; Elsevier Academic Press: Burlington, MA, 2005.
(13) Supovitz, J. A.; Turner, H. M. The Effects of Professional Development on Science Teaching and Practices in Classroom Culture. *J. Res. Sci. Teach.* **2000**, *37* (9), 963−980.
(14) Garet, M. S.; Porter, A. C.; Desimone, L.; Birman, B. F.; Yoon, K. S. What Makes Professional Development Effective? Results from a National Sample of Teachers. *Am. Educ. Res. J.* **2001**, *38* (4), 915−945.
(15) RISE Programs Web page. http://www.rise.duke.edu/apep (accessed Jan 2014).
(16) National Research Council. National Committee on Science Education Standards and Assessment. *National Science Education Standards*; The National Academies Press: Washington, DC, 1996.
(17) Next Generation Science Standards Home page. http://www.nextgenscience.org (accessed Jan 2014).
(18) O'Sullivan, C. Y.; Reese, C. M.; Mazzeo, J. *NAEP 1996 Science Report Card for the Nation and the States*; National Center for Education Statistics: Washington, DC, 1997.
(19) Shrout, P. E. Measurement Reliability and Agreement in Psychiatry. *Stat. Methods Med. Res.* **1998**, *3*, 301−317.
(20) Raudenbush, S. W.; Bryk, A. S. *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd ed.; Sage Publications: New York, 2002.
(21) Gelman, A.; Carlin, J. B.; Stern, H. S.; Rubin, D. B. *Bayesian Data Analysis*, 2nd ed.; Chapman and Hall: Boca Raton, FL, 2004.
(22) Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Earlbaum Associates: Hillsdale, NJ, 1988.
(23) National Research Council. Committee on a Conceptual Framework for New K-12 Science Education Standards, Board on Science Education, Division of Behavioral and Social Sciences and Education. In *A Framework for K−12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; The National Academies Press: Washington, DC, 2012.
(24) Kennedy, M. M. *Form and Substance in In-Service Teacher Education*, Research Monograph No. 13; National Institute for Science Education: Madison, WI, 1998.
(25) Furtak, E. M.; Seidel, T.; Iverson, H.; Briggs, D. C. Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis. *Rev. Educ. Res.* **2012**, *82* (3), 300−329.
(26) Posner, G. J.; Strike, K. A.; Hewson, P. W.; Gertzog, W. A. Accommodation of a Scientific Conception: Toward a Theory of Conceptual Change. *Sci. Educ.* **1982**, *66*, 211−227.

SUPPORTING INFORMATION

Supplementary Table S1. Demographics of teachers participating in APEP

Category	Conference-based (N = 100)	Distance Learning (N = 56)
Subject		
Biology	36	28
Chemistry	40	21
Biology & Chemistry	21	10
Other	3	4
School Type		
Public	84	53
Private	13	2
Parochial	3	1
School Size		
< 500	19	19
501-1000	28	9
1001-2000	36	10
>2000	11	18
School Minority Population		
< 10%	30	14
10-39%	41	25
40-79%	16	15
> 80%	13	2
School Locale		
Urban	30	5
Suburban	46	26
Rural	24	25
US Location		
Northeast	9	0
Southeast	39	20
Midwest	49	6
West	2	30
Military	1	0

Supplementary Table S2. Number of Teachers Participating in the Distance Learning Workshops as a Function of Geographic Location

Location	Number of Teachers (N=56)
Southeast	
Coral Glades, FL	3
Lugoff, SC**	6
Memphis, TN	6
Abingdon, VA	5
Midwest	
Thornton, CO*	2
Houston, TX	3
Fort Worth, TX†	15
Johnston, IA*	4
Kansas City, MO*	2
West	
Concord, CA	3
Hayward, CA	3
Pine Valley, CA	1
Bellingham, WA	3

*Streaming participants; **included 2 broadcast sites; †included 5 broadcast sites

Supplementary Table S3: Teacher Evaluation of APEP Workshops

	Distance Learning Workshop		Conference-based Workshop	
Evaluation Item (Strongly Agree = 5; Strongly Disagree = 1)	Mean Score (N=51)	% Answering 4 or 5	Mean Score (N=99)	% Answering 4 or 5
I learned something new in chemistry.	4.3	84.3	4.4	85.9
I learned something new in biology.	4.6	94.1	4.7	94.9
I learned something new in pharmacology.	4.8	98.0	4.8	97.0
I learned something new about integrating biology & chemistry through pharmacology.	4.7	98.0	4.8	100.0
This workshop stimulated my thinking about new ways of teaching my subject.	4.7	100	4.9	100.0
This workshop stimulated my thinking about new ways to integrate chemistry and biology.	4.6	96.1	4.8	100.0
I found the DL approach effective.	4.2	86.3	N/A	N/A
I had no trouble following the lectures delivered via TV.	4.0	80.4	N/A	N/A
I prefer the DL approach to traveling overnight to a workshop.	4.0	76.5	N/A	N/A
I prefer traveling to a workshop versus a distance learning approach.	N/A	N/A	4.8	99.0

Example of the Story (the Problem)
Module 4 – Alcohol and the Breathalyzer Test

Spritney Beers Blows it Bigtime

The big news is that celebrity singer, Spritney Beers has been arrested for driving under the influence (DUI). Spritney blew a "0.08%" on the Breathalyzer™ when she was stopped by the police after driving over the center line. This value reflects the blood alcohol concentration (BAC) and, in California, it is the value that legally designates intoxication. Despite considerable evidence that the Breathalyzer™ test gives an accurate approximation of the BAC, there is still much discussion about the validity of this test. While in jail, Spritney used her one phone call to contact her attorney.

A court date has been set for three weeks from today. Paparazzi photographed her as she left the courthouse in a heated discussion with her attorney. Reporters at the scene quoted him saying "I am very suspect of the validity of the Breathalyzer™ test given to my client. I only have to establish reasonable doubt in this case."

Back at school, you are discussing the case with a classmate when your science teacher interrupts and decides that the topic is worthy of a class debate. To prepare for the class debate, you decide to learn how a Breathalyzer™ test is used to determine the BAC so that you can debate whether it is accurate or not.

Spritney Beers' level of alcohol intoxication was assessed by measuring the concentration of ethanol she expelled from her lungs. This technique works because a water-based molecule like ethanol that is absorbed from the gut into the bloodstream will reach the lungs, where it is exhaled as a vapor in the air.

1. Diagram the path that alcohol will take from the gut (stomach & small intestine) to the lungs via the circulatory system. Which membranes will ethanol have to cross? Although some alcohol is exhaled from the lungs, the rest stays in the blood. Where does the ethanol in the blood go once it leaves the lungs?

Ethanol is dissolved in the blood and is distributed to organs around the body. It is a **volatile** molecule and can be vaporized quite easily. In the lung, ethanol is converted from a liquid to a gas, so it can be exhaled in the air.

2. Identify and describe the chemical and physical properties of ethanol that contribute to its volatility. It will help to draw the chemical structure of ethanol.

3. Where in the lung is ethanol vaporized? What role does the lung play in the **vaporization** of ethanol?

The Breathalyzer™ can approximate a person's blood alcohol concentration (BAC) because the concentration of alcohol vapor in the lungs is directly related to its concentration in the blood. Alcohol vaporizes in the air sacs (alveoli) of the lungs and achieves an **equilibrium** with the concentration of alcohol that is still in the blood.

4. What is meant by equilibrium? Are ethanol molecules still moving across the membrane between the capillary and the alveolus?

5. What would happen to the equilibrium if some of the ethanol leaves the alveolar sac by exhalation?

When a person exhales into a breath analyzer such as the Breathalyzer™ tube, the exhaled alcohol reacts with compounds in the Breathalyzer™ chamber to produce a change in color from orange to green. The chemical reaction indicated by the color change involves **oxidation** and **reduction**. Silver nitrate catalyzes the reaction.

6. What is **oxidized** in the Breathalyzer™? What is **reduced**? What role does the **catalyst** play in the chemical reaction?

The degree of the color change indicates how much alcohol is present in the expired air and the instrument calculates an actual concentration. However, for the Breathalyzer™ to calculate how much alcohol is present in the expired air sample it must take into account the volume of blood from which the alcohol originated. There is a standard way of describing this relationship; it is called the blood-to-breath ratio or **partition ratio**. The average blood-to-breath ratio is 2100:1 and this is the value used for legal purposes. The ratio assumes that an equilibrium exists between the blood and the alveolar air.

7. Explain what this ratio means in terms of the concentration of alcohol in the blood and the breath. Does the ratio change as one exhales?

Actually, the ratio can vary between 1500:1 and 3000:1 depending upon a number of factors including a person's age, gender, and genetic makeup. In Spritny Beers' case, the reported BAC was 0.08%, based on the 2100:1 ratio. In fact, she may be anywhere in this range from 1500:1 to 3000:1.

8. Calculate the underestimation and overestimation of Spritney's BAC assuming she had a blood-to-breath ratio of 1500:1 and 3000:1. Would Spritney still be considered legally intoxicated based on your answers?

As a point of discussion, consider the advantages and disadvantages of using the Breathalyzer™ test and decide for yourself whether you would prosecute based on the Breathalyzer™ evidence in this case. Is there other evidence to consider?

Sample Teacher Assessment Items

(Correct answers are underlined.)

Chemistry-related items:

When hydrogen is burned it combines with oxygen to produce water molecules. Which describes the state of each atom?

A. <u>Hydrogen is oxidized, Oxygen is reduced</u>
B. Oxygen is oxidized, Hydrogen is reduced
C. Both atoms are oxidized
D. Both atoms are reduced

Ethanol, an alcohol, is a molecule that has both a polar end (OH group) and a non-polar end containing a chain of 2 C atoms. Which of the following describes the solubility of alcohol molecules with longer C atom chains?

A. they become more soluble in water because they are more polar
B. they become more soluble in water because they are more non-polar
C. they become less soluble in water because they are more polar
D. <u>they become less soluble in water because they are more non-polar</u>

Biology-related items:

Why do neurons use both electrical and chemical signals to communicate with each other?

A. Neurons need chemical signals to stop the electrical signals
B. Neurons need fast and slow ways to pass information between them
C. <u>Electrical signals can't jump across the synaptic space</u>
D. Electrical signals can degrade over the distance of an axon

A molecule is in the capillaries. What path does it take to move through the circulation?

A. <u>veins, right side of the heart, lungs, left side of the heart, arteries</u>
B. veins, left side of the heart, lungs, right side of the heart, arteries
C. arteries, right side of the heart, lungs, left side of the heart, veins
D. arteries, left side of the heart, lungs, right side of the heart, veins

Sample Student Assessment Items

(Correct answers are underlined.)

Basic knowledge assessment of chemistry

What feature of the hydroxyl group (–OH) makes it possible to form hydrogen bonds with other molecules?

A. electronegative oxygen contains unshared pairs of electrons
B. electronegative oxygen contains an extra electron
C. electropositive oxygen contains unshared pairs of electrons
D. electropositive oxygen contains an extra electron
E. don't know

Methanol is more volatile than water because:

A. it has a higher boiling point
B. it has a lower boiling point
C. it can form more covalent bonds
D. it can form more hydrogen bonds
E. don't know

A chain of carbon atoms single-bonded together and saturated with hydrogen atoms (-CH2-) produces a non-polar molecule because:

A. the electron pairs between the carbons are shared unequally
B. the electron pairs between the carbons are shared equally
C. there is an unpaired electron between the carbons
D. there is an equal number of electrons at each end of the molecule
E. don't know

Basic knowledge assessment of biology

Protein synthesis proceeds in the following order:

A. DNA is translated to mRNA, mRNA carries out transcription to a protein
B. DNA is transcribed to mRNA, mRNA carries out translation to a protein
C. RNA is transcribed to DNA, DNA carries out translation to a protein
D. RNA is translated to DNA, DNA carries out transcription to a protein
E. don't know

The ability of an electrical impulse to flow along neurons is based on the distribution of ions inside and outside the cell. What is the distribution of ions when the neuron is at rest?

A. K+ high on the outside, Na+ high on the inside
B. K+ high on the outside, Ca++ high on the inside
C. Na+ high on the outside, K+ high on the inside
D. Na+ high on the outside, Ca++ high on the inside
E. don't know

The brain depends on having enough oxygen to work properly. Trace the path that the oxygen takes to get up to the brain once you breathe it in:

A. lungs, right side of the heart, arteries, brain
B. lungs, left side of the heart, arteries, brain
C. lungs, right side of the heart, veins, brain
D. lungs, left side of the heart, veins, brain
E. don’t know

Advanced knowledge assessment of chemistry

Why would a police officer ask a person stopped for driving while impaired to take 3 deep breaths before blowing into the Breathalyzer?

A. to increase the alcohol vapor in the mouth enough to get a reading on the Breathalyzer
B. to get rid of alcohol vapor in the stomach so it won’t confound the measurement
C. to accurately reflect the equilibrium concentration of alcohol in the mouth with respect to that in the blood
D. to accurately reflect the equilibrium concentration of alcohol in the lungs with respect to that in the blood
E. don’t know

Advanced knowledge assessment of biology

Alcohol causes many effects such as poor judgment, drowsiness, and incoordination. Which brain areas are affected to produce these problems (in order)?

A. cerebellum, limbic system, hypothalamus
B. frontal cortex, brainstem, cerebellum
C. limbic system, cerebellum, frontal cortex
D. hypothalamus, frontal cortex, limbic system
E. don’t know

Statistical Methods

We use Bayesian methods to estimate the models via the WinBUGS software package (Spiegelhalter et al., 2003). Bayesian methods start with statements of prior beliefs about the model parameters. The prior beliefs are combined with the observed data to update knowledge about the model parameters, resulting in a posterior distribution for the parameters. We assume vague prior beliefs so that the posterior distribution is dominated by the observed data. We fit the models with alternate prior assumptions and found that the results are insensitive to reasonable specifications of prior beliefs.

We estimate the effects of the modules using binomial logistic regressions with random effects for teachers. These models enable us to control for demographic variables that could influence test performance. They also account for the fact that students taught by the same teacher are more likely to have similar test scores than two students taught by different teachers; that is, the models control for teacher effects. We use separate models for the basic knowledge and advanced knowledge tests.

Formally, we model the number of correct answers Y_{ij} of student i in the class of teacher j as Y_{ij}~binomial(n, p_{ij}), where

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = X_{ij}\,\beta + m_{ij1}b_1 + m_{ij2}b_2 + m_{ij3}b_3 + m_{ij4}b_4 + \lambda_{ij}.$$

Here, n refers to the number of questions in the test under consideration. X_{ij} is a vector of student attributes including race, gender, current course, and two indicator variables: the first indicator equals one for students in the experimental condition and whose teacher attended the distance learning workshop (and the indicator equals zero otherwise), and the second indicator equals one for students in the experimental condition and whose teacher attended the conference-based workshop (and the indicator equals zero otherwise). β is a vector of regression coefficients for the students' attributes. The m_{ij} terms are indicator variables for the number of modules used in the student's class. For example, if two modules are used, we have $m_{ij2} = 1$ and $m_{ij1} = m_{ij3} = m_{ij4} = 0$; the corresponding b coefficients indicate the effect associated with using that number of modules. We use four indicator variables instead of one continuous variable to allow the response to vary flexibly with module use.

The λ_{ij} random effects are interpreted as student-specific ability; for example, $\lambda_{ij} > 0$ means that the student performs better than the average student with the same array of covariates. We assume that these random effects follow the normal distribution

λ_{ij}~normal$(\alpha_j, \sigma_\lambda^2)$. Finally, we assume that the teacher-level random effects follow the normal distribution α_j~normal$(0, \tau^2)$.

We restrict the sample to include students whose teachers were in both the experimental and control conditions. Because teachers are present in both conditions, this eliminates differences in teacher characteristics across the conditions and hence reduces the potential for confounding. The conclusions do not change substantially when we expand the analysis to include teachers with only control (2877 students) or only experimental scores (147 students).

Finally, we assessed the fit of our models in two ways. First, we fit a more complex model that relaxes the assumption that the questions are of equal difficulty. This did not improve the fit according to the DIC, which is a model selection criterion for hierarchical models (Spiegelhalter et al., 2002). Second, for the models used to construct the tables, we performed posterior predictive checks (Gelman et al., 2004). This involves simulating new outcomes for the students in the dataset based on the fitted model and then comparing the simulated outcomes to the true Y_{ij}. Discrepancies in the simulated and actual outcomes indicate inadequate fit. These predictive checks do not reveal problems with the model specification.

Supplementary Table S4: Basic knowledge regression.

Predictor	Mean	Standard Error	95% interval	Mean Probability of Correct Answer (95% interval)
Intercept	-1.426	0.032	(-1.491, -1.365)	0.19 (0.18 , 0.20)
Male	**0.098**	**0.012**	**(0.074, 0.121)**	**0.21 (0.20, 0.22)**
Asian	**0.100**	**0.025**	**(0.052, 0.148)**	**0.21 (0.20, 0.22)**
Black	**-0.080**	**0.022**	**(-0.123, -0.037)**	**0.18 (0.17, 0.19)**
Hispanic	-0.032	0.024	(-0.078, 0.014)	0.19 (0.18, 0.20)
Native American	0.069	0.042	(-0.014, 0.149)	0.20 (0.19, 0.22)
Chemistry 1	**0.178**	**0.030**	**(0.119, 0.236)**	**0.22 (0.21, 0.23)**
Chemistry 2	**0.611**	**0.043**	**(0.527, 0.693)**	**0.31 (0.29, 0.33)**
Biology 2	**0.460**	**0.026**	**(0.409, 0.511)**	**0.28 (0.26, 0.29)**
Conference workshop	**0.090**	**0.023**	**(0.044, 0.136)**	**0.21 (0.20, 0.22)**
Distance workshop	0.034	0.028	(-0.022, 0.088)	0.20 (0.19, 0.21)
1 Module	**0.084**	**0.031**	**(0.023, 0.146)**	**0.22 (0.21, 0.24)**
2 Modules	**0.118**	**0.032**	**(0.054, 0.180)**	**0.23 (0.21, 0.24)**
3 Modules	**0.165**	**0.036**	**(0.093, 0.235)**	**0.24 (0.22, 0.25)**
4 Modules	**0.197**	**0.030**	**(0.137, 0.256)**	**0.24 (0.23, 0.26)**

Supplementary Table S5: Advanced knowledge regression.

Predictor	Mean	Standard Error	95% interval	Mean Probability of Correct Answer (95% interval)
Intercept	-1.138	0.038	(-1.210, -1.064)	0.24 (0.23, 0.26)
Male	**-0.054**	**0.017**	**(-0.087, -0.020)**	**0.23 (0.22, 0.25)**
Asian	0.024	0.035	(-0.044, 0.094)	0.25 (0.23, 0.26)
Black	**-0.186**	**0.031**	**(-0.246, -0.127)**	**0.21 (0.20, 0.23)**
Hispanic	**-0.079**	**0.035**	**(-0.147, -0.012)**	**0.23 (0.21, 0.25)**
Native American	-0.010	0.061	(-0.124, 0.109)	0.24 (0.22, 0.27)
Chemistry 1	0.070	0.040	(-0.006, 0.152)	0.26 (0.24, 0.27)
Chemistry 2	**0.419**	**0.060**	**(0.309, 0.537)**	**0.33 (0.30, 0.35)**
Biology 2	**0.404**	**0.038**	**(0.331, 0.477)**	**0.32 (0.31, 0.34)**
Conference workshop	0.042	0.034	(-0.023, 0.109)	0.25 (0.23, 0.27)
Distance workshop	**-0.121**	**0.041**	**(-0.201, -0.039)**	**0.22 (0.20, 0.24)**
1 Module	**0.186**	**0.046**	**(0.096, 0.277)**	**0.29 (0.27, 0.31)**
2 Modules	**0.200**	**0.046**	**(0.110, 0.289)**	**0.29 (0.27, 0.31)**
3 Modules	**0.348**	**0.052**	**(0.249, 0.448)**	**0.32 (0.30, 0.35)**
4 Modules	**0.420**	**0.043**	**(0.335, 0.504)**	**0.34 (0.32, 0.36)**

For both regressions, the base condition (i.e., the intercept) is a white female in a Biology 1 class whose teacher did not yet attend the workshop and did not use any modules. Probabilities in other rows change one of those characteristics, except for the module use rows, which also assume that the teacher attended the conference-based workshop. Coefficients of variables with 95% intervals that do not contain zero are significantly different from zero (bolded). Interpretations of the coefficients as odds ratios are obtained by exponentiation of the point estimates and confidence interval limits. For example, a Biology 2 student has $e^{.404} = 1.50$ times the odds of answering an advanced knowledge question correctly compared to a Biology 1 student. The 95% interval for the odds ratio is $e^{.331} = 1.39$ to $e^{.447} = 1.56$.

Gelman A, Carlin JB, Stern HS, and Rubin DB (2004) *Bayesian Data Analysis*, 2nd ed, Chapman and Hall: Boca Raton, FL.

Spiegelhalter DJ, Best NG, Carlin BP, and van der Linde A (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B.* (64)4:583-639.

Spiegelhalter DJ, Thomas A, Best NG, Gilks WR, and Lunn D (2003) BUGS: Bayesian inference using Gibbs sampling. MRC Biostatistics Unit, Cambridge, England. www.mrc-bsu.cam.ac.uk/bugs